

**ĐẠI HỌC QUỐC GIA HÀ NỘI  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ**

---

**LÊ THỊ KIM NGA**

**NGHIÊN CỨU PHÁT HIỆN  
MẪU CHẤT LIỆU TRONG ẢNH**

**LUẬN ÁN TIẾN SĨ CÔNG NGHỆ THÔNG TIN**

**HÀ NỘI – 2014**

**ĐẠI HỌC QUỐC GIA HÀ NỘI  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ**

---

**LÊ THỊ KIM NGÀ**

**NGHIÊN CỨU PHÁT HIỆN  
MẪU CHẤT LIỆU TRONG ẢNH**

Chuyên ngành: Khoa học máy tính

Mã số: 62 48 01 01

**LUẬN ÁN TIẾN SĨ CÔNG NGHỆ THÔNG TIN**

NGƯỜI HƯỚNG DẪN KHOA HỌC:

1. PGS.TS. ĐỖ NĂNG TOÀN
2. PGS.TS. ĐINH MẠNH TƯỜNG

**HÀ NỘI - 2014**

## **Lời cam đoan**

Tôi xin cam đoan đây là công trình nghiên cứu của riêng tôi. Các kết quả được viết chung với các tác giả khác đều được sự đồng ý của đồng tác giả trước khi đưa vào luận án. Các kết quả nêu trong luận án là trung thực và chưa từng được ai công bố trong các công trình nào khác.

Tác giả

Lê Thị Kim Nga

## **Lời cảm ơn**

Luận án được thực hiện tại Trường Đại học Công nghệ - Đại học Quốc Gia Hà Nội và Viện Công nghệ thông tin - Viện Hàn lâm Khoa học và Công nghệ Việt Nam, dưới sự hướng dẫn của PGS.TS. Đỗ Năng Toàn và PGS.TS. Đinh Mạnh Tường.

Tôi xin bày tỏ lòng biết ơn sâu sắc đến PGS.TS. Đỗ Năng Toàn và PGS.TS. Đinh Mạnh Tường, các Thầy đã có những định hướng giúp tôi thành công trong công việc nghiên cứu của mình. Thầy cũng động viên chỉ bảo cho tôi vượt qua những khó khăn và cho tôi nhiều kiến thức quý báu về nghiên cứu khoa học. Nhờ sự chỉ bảo của Thầy, tôi mới có thể hoàn thành luận án.

Tôi vô cùng cảm ơn PGS.TS. Hoàng Xuân Huân và GS.TS. Nguyễn Thanh Thủy, các Thầy đã nhiệt tình giúp đỡ tôi rất nhiều trong suốt quá trình học tập, nghiên cứu và hiệu chỉnh luận án.

Tôi xin gửi lời cảm ơn sâu sắc đến PGS.TS. Đỗ Trung Tuấn, PGS. TS. Bùi Thế Duy, PGS.TS. Trịnh Nhật Tiến, TS. Nguyễn Văn Vinh và TS. Nguyễn Ngọc Hóa, các Thầy đã giúp tôi rất nhiều trong việc hoàn thiện luận án.

Tôi xin chân thành cảm ơn các Thầy, Cô Khoa Công nghệ thông tin, Trường Đại học Công nghệ, Đại Học Quốc Gia Hà Nội, các anh chị em cán bộ trong phòng Công nghệ Thực tại ảo, Viện Công nghệ thông tin đã tạo mọi điều kiện thuận lợi và đóng góp ý kiến cho tôi trong quá trình làm nghiên cứu sinh.

Đặc biệt tôi xin chân thành cảm ơn lãnh đạo Nhà trường và Khoa Công nghệ thông tin, Trường Đại học Quy Nhơn đã tạo điều kiện thuận lợi và hỗ trợ tôi trong quá trình học tập và làm luận án.

Cuối cùng, tôi xin gửi lời cảm ơn sâu sắc đến gia đình và bạn bè, đã tạo cho tôi điểm tựa vững chắc để có được thành công như hôm nay.

# MỤC LỤC

Lời cam đoan .....	1
Lời cảm ơn .....	2
MỤC LỤC .....	3
Danh mục các ký hiệu và chữ viết tắt.....	7
Danh mục các bảng .....	9
Danh mục các hình vẽ, đồ thị.....	10
MỞ ĐẦU .....	13
Chương 1. TỔNG QUAN VỀ PHÁT HIỆN MẪU CHẤT LIỆU TRONG ẢNH .....	23
1.1. Chất liệu và bài toán phát hiện mẫu chất liệu trong ảnh .....	23
1.1.1. Chất liệu và mẫu chất liệu trong ảnh .....	23
1.1.2. Bài toán phát hiện mẫu chất liệu trong ảnh .....	26
1.1.3. Các thách thức của phát hiện mẫu chất liệu trong ảnh .....	27
1.2. Các cách tiếp cận phát hiện mẫu chất liệu trong ảnh .....	29
1.2.1. Tiếp cận dựa vào đặc trưng địa phương .....	30
1.2.1.1. Phương pháp dựa trên độ cong của đường biên.....	31
1.2.1.2. Phương pháp dựa trên cường độ ảnh .....	31
1.2.1.3. Phương pháp định hướng bất biến với các phép biến đổi.....	32
1.2.1.4. Phương pháp tỉ lệ chu vi và diện tích.....	32
1.2.1.5. Phương pháp cấu trúc hình học.....	33
1.2.2. Tiếp cận dựa vào đặc trưng toàn cục .....	33
1.2.2.1. Phương pháp lược đồ màu .....	34
1.2.2.2. Phương pháp ma trận đồng hiện mức xám .....	34

1.2.2.3. Phương pháp mẫu nhị phân địa phương .....	34
1.2.2.4. Phương pháp dựa vào nhiều chất liệu .....	35
1.2.3. Tiếp cận dựa vào mô hình .....	39
1.2.3.1. Mô hình SAR và RISAR.....	39
1.2.3.2. Mô hình Markov .....	40
1.2.3.3. Mô hình hình học Fractal .....	41
1.3. Kết luận và vấn đề nghiên cứu .....	47
<b>Chương 2. PHÁT HIỆN MẪU CHẤT LIỆU DỰA VÀO ĐẶC TRƯNG BẤT BIẾN</b> <b>ĐỊA PHƯƠNG .....</b>	<b>48</b>
2.1. Đặt vấn đề.....	48
2.2. Trích chọn đặc trưng bất biến địa phương cho mẫu chất liệu .....	49
2.2.1. Tìm các điểm bất biến địa phương trên không gian tỉ lệ .....	49
2.2.2. Xây dựng mô tả địa phương .....	53
2.3. Phát hiện mẫu chất liệu dựa vào đặc trưng bất biến địa phương .....	54
2.3.1. Đối sánh dựa vào phương pháp lân cận gần nhất .....	55
2.3.2. Xác định sự tương ứng của mẫu chất liệu trong ảnh .....	55
2.3.3. Phân cụm các điểm ứng cử trung tâm .....	56
2.3.4. Thuật toán phát hiện mẫu chất liệu DMBLIF.....	57
2.3.5. Thực nghiệm.....	61
2.4. Phát hiện ảnh số giả mạo dựa vào thuật toán DMBLIF .....	65
2.4.1. Ảnh số giả mạo và các dạng ảnh số giả mạo cơ bản .....	65
2.4.1.1. Ảnh số giả mạo.....	66
2.4.1.2. Các dạng ảnh số giả mạo cơ bản.....	67

2.4.2. Thuật toán phát hiện ảnh số giả mạo KPFIImage.....	69
2.4.2.1. Thuật toán phát hiện ảnh số giả mạo Exact Match .....	72
2.4.2.2. Thuật toán KPFIImage .....	74
2.4.3. Thực nghiệm.....	78
2.5. Kết luận chương 2 .....	83
Chương 3. PHÁT HIỆN MẪU CHẤT LIỆU DỰA VÀO ĐẶC TRƯNG NHIỄU .....	84
3.1. Biểu diễn mẫu chất liệu dựa đặc trưng nhiễu .....	84
3.1.1. Đặc trưng nhiễu chất liệu.....	85
3.1.2. Xây dựng đặc trưng nhiễu cho mô hình mẫu chất liệu.....	85
3.1.3. Thuật toán biểu diễn mẫu chất liệu dựa vào đặc trưng nhiễu RMBN ...	87
3.2. Phát hiện mẫu chất liệu dựa vào đặc trưng mô hình nhiễu chất liệu.....	90
3.2.1. Phân lớp mẫu chất liệu dựa vào hàm phân phối Gauss .....	90
3.2.2. Thuật toán phát hiện mẫu chất liệu dựa vào nhiễu DMBNF.....	91
3.2.3. Thực nghiệm.....	94
3.3. Kết luận chương 3 .....	97
Chương 4. PHÁT HIỆN MẪU CHẤT LIỆU DỰA VÀO HÌNH HỌC FRACTAL ...	99
4.1. Đặt vấn đề.....	99
4.2. Cơ sở toán học .....	101
4.3. Biểu diễn mẫu chất liệu dựa vào hình học Fractal .....	103
4.3.1. Đặc trưng hình học Fractal cho chất liệu.....	103
4.3.2. Xây dựng mô tả Fractal cho mẫu chất liệu .....	104
4.3.3. Thuật toán biểu diễn mẫu chất liệu dựa vào Fractal RMBF.....	107
4.4. Thuật toán phát hiện mẫu chất liệu dựa vào Fractal DMBF .....	110

4.5. Kết luận chương 4 .....	112
KẾT LUẬN .....	113
DANH MỤC CÁC CÔNG TRÌNH KHOA HỌC CỦA TÁC GIẢ LIÊN QUAN ĐẾN LUẬN ÁN .....	115
TÀI LIỆU THAM KHẢO .....	116
PHỤ LỤC .....	128



## **Danh mục các ký hiệu và chữ viết tắt**

BRDF	Bidirectional Reflectance Distribution Function (Hàm phân phối phản xạ hai chiều)
BTF	Bidirectional Texture Function (Hàm texture hai chiều)
CBIR	Content Based Image Retrieval (Tra cứu ảnh dựa trên nội dung)
CSAR	Circular Simultaneous Autoregressive Model (Mô hình tự hồi quy đồng tâm)
DMBLIF	Thuật toán phát hiện mẫu chất liệu dựa vào đặc trưng bất biến địa phương
DMBNF	Thuật toán phát hiện mẫu chất liệu dựa vào đặc trưng nhiễu
DMBF	Thuật toán phát hiện mẫu chất liệu dựa vào Fractal
GLCM	Grey Level Co-occurrence Matrix (Ma trận đồng hiện mức xám)
HMM	Hidden Markov Model (Mô hình Markov ẩn)
IFS	Iterated Function System (Hệ hàm lặp)
ISODATA	Iterative Self-Organizing Data Analysis Technique (Thuật toán phân cụm)
Keypoint	Điểm bất biến tỉ lệ
KPFImage	Thuật toán phát hiện ảnh giả mạo dựa vào tiếp cận phát hiện chất liệu

LBP	Local Binary Pattern (Mẫu nhị phân địa phương)
MRF	Markov Random Field (Trường Markov ngẫu nhiên)
PCA	Principal Component Analysis (Phân tích thành phần chính)
$Rec(\mathfrak{R}_i)$	Khối bao (vị trí và kích thước) của vùng ảnh con $\mathfrak{R}_i$
RMBF	Thuật toán biểu diễn mẫu chất liệu dựa vào Fractal
RMBN	Thuật toán mô tả mẫu chất liệu dựa vào đặc trưng nhiễu
RISAR	Rotation Invariant Simultaneous AutoRegressive Model (Mô hình tự hồi quy đồng thời bất biến quay)
SAR	Simultaneous AutoRegressive Model (Mô hình tự hồi quy đồng thời)
Scale	Tỉ lệ
SIFT	Scale Invariant Feature Transform (Biến đổi đặc trưng bất biến tỉ lệ)
Texture	Kết cấu bề mặt

## **Danh mục các bảng**

Bảng 2.1. Kết quả đánh giá thuật toán DMBLIF.....	62
Bảng 3.1. Kết quả đánh giá thuật toán DMBNF.....	94

## Danh mục các hình vẽ, đồ thị

Hình 1.1. Sự phản chiếu theo một hướng (Specular Reflection).....	24
Hình 1.2. Sự phản chiếu khuếch tán (Diffuse Reflection).....	25
Hình 1.3. Một số mẫu chất liệu thông thường .....	26
Hình 1.4. Sơ đồ hệ thống phát hiện mẫu chất liệu tổng quát.....	27
Hình 1.5. Các thể hiện khác nhau của cùng một mẫu chất liệu.....	28
Hình 1.6. Các phép biến đổi hình học trong quá trình thu nhận ảnh.....	29
Hình 1.7. Nhiều mẫu chất liệu, dòng trên: các ảnh gốc; dòng bên dưới: các ảnh nhiều tương ứng .....	36
Hình 1.8. Sơ đồ phân loại nhiều .....	38
Hình 2.1. $D(x,y,\sigma)$ xấp xỉ với $L(x,y,\sigma)$ .....	51
Hình 2.2. Xây dựng mô tả cho điểm bất biến tỉ lệ.....	54
Hình 2.3. Cách xác định vec tơ định vị .....	56
Hình 2.4. Kết quả phát hiện mẫu chất liệu cỏ bằng thuật toán DMBLIF (a) Ảnh vào; (b) Mẫu chất liệu cỏ ; (c) Kết quả phát hiện mẫu (b); (d) Mẫu cỏ bị quay; (e) Kết quả phát hiện mẫu (d); (f) Mẫu cỏ thay đổi quay và tỉ lệ; (g) Kết quả phát hiện mẫu (f).....	63
Hình 2.5. Kết quả phát hiện mẫu chất liệu gỗ bằng thuật toán DMBLIF (a) Ảnh vào; (b) Mẫu gỗ bị thay đổi tỉ lệ; (c) Kết quả phát hiện khi mẫu gỗ thay đổi quay và tỉ lệ ; (d) Mẫu gỗ bị quay ; (d) Kết quả phát hiện mẫu gỗ đã bị thay đổi quay ; (e) Mẫu gỗ bị thay đổi tỉ lệ và quay ; (f) Kết quả phát hiện mẫu gỗ bị thay đổi quay và thay đổi tỉ lệ. ....	64
Hình 2.6. Minh họa về việc giả mạo ảnh.....	66

Hình 2.7. Ghép ảnh từ 2 ảnh riêng rẽ (a) Ảnh ghép từ hai ảnh riêng rẽ; (b) Ảnh ghép từ hai ảnh có thay đổi tỉ lệ .....	68
Hình 2.8. Ảnh che phủ và bỏ đi đối tượng (a), (c) Ảnh gốc; (b), (d) Ảnh đã che phủ đối tượng; .....	69
Hình 2.9. Ảnh bổ sung đối tượng (a) Ảnh gốc; (b) Ảnh bổ sung đối tượng .....	69
Hình 2.10. Ảnh giả mạo cắt dán bởi bổ sung đối tượng (a) Ảnh gốc; (b) Ảnh giả mạo bổ sung đối tượng .....	71
Hình 2.11. Tìm kiếm khối bao của thuật toán Exact Match .....	73
Hình 2.12. Kết quả phát hiện giả mạo bằng thuật toán Exact match (a) Ảnh gốc; (b) Ảnh giả mạo cắt dán; (c) Các vùng giả mạo được phát hiện bởi thuật toán Exact Match.....	73
Hình 2.13. Kết quả phát hiện ảnh giả mạo cắt dán với ảnh 24 bit màu (a) Ảnh gốc; (b) Ảnh giả; (c) Phát hiện giả mạo bằng Exact match; (d) Phát hiện giả mạo bằng Exact match*; (e) Phát hiện giả mạo bằng KPFIImage .....	80
Hình 2.14. Kết quả phát hiện ảnh giả mạo cắt dán với ảnh đa mức xám (a) Ảnh gốc; (b) Ảnh giả; (c) Phát hiện giả mạo bằng Exact match; (d) Phát hiện giả mạo bằng Exact match*; (e) Phát hiện giả mạo bằng KPFIImage .	81
Hình 2.15. Kết quả phát hiện ảnh giả mạo dạng nén (a) Ảnh gốc; (b) Ảnh giả với bông hoa bị thay đổi quay; (c) Kết quả phát hiện bằng Exact match* của (b); (d) Ảnh giả với bông hoa bị thay đổi tỉ lệ và quay; (e) Kết quả phát hiện bằng KPFIImage của (d).....	82
Hình 3.1. Mô hình xây dựng đặc trưng nhiễu cho mẫu chất liệu .....	86
Hình 3.2. Kết quả tìm đặc trưng nhiễu của mẫu gỗ 1 của thuật toán RMBN (a) Mẫu chất liệu gỗ 1 dưới dạng mô hình gồm tập các chất liệu gỗ; (b) Đặc trưng nhiễu của mẫu chất liệu gỗ 1 .....	89

Hình 3.3. Kết quả tìm đặc trưng nhiễu của mẫu gỗ 2 của thuật toán RMBN (a) Mẫu chất liệu gỗ 2 dưới dạng mô hình gồm tập các chất liệu gỗ; (b) Đặc trưng nhiễu của mẫu gỗ 2.....	89
Hình 3.4. Kết quả tìm đặc trưng nhiễu của mẫu gỗ 3 của thuật toán RMBN (a) Mẫu chất liệu gỗ 3 dưới dạng mô hình gồm tập các chất liệu gỗ; (b) Đặc trưng nhiễu của mẫu chất liệu gỗ 3 .....	89
Hình 3.5. Minh họa vùng chất liệu R được chọn dựa vào phân phối Gauss .....	91
Hình 3.6. Các mẫu nhiễu của một số chất liệu (a) Mẫu chất liệu gỗ; (b) Đặc trưng nhiễu của mẫu gỗ; (c) Mẫu chất liệu cỏ; (d) Đặc trưng nhiễu của mẫu cỏ .....	95
Hình 3.7. Kết quả phát hiện mẫu chất liệu cỏ cây bằng thuật toán DMBNF (a) Mẫu chất liệu cỏ cây; (b) Ảnh vào; (c) Đặc trưng nhiễu của chất liệu cỏ cây; (d) Ảnh nhiễu của ảnh vào (e) Kết quả phát hiện trên ảnh nhiễu vào; (f) Kết quả phát hiện trên ảnh vào.....	96
Hình 3.8. Kết quả phát hiện mẫu chất liệu gỗ bằng thuật toán DMBNF (a) Mẫu chất liệu gỗ; (b) Ảnh vào; (c) Đặc trưng nhiễu của mẫu chất liệu gỗ; (d) Nhiễu của ảnh vào; (e) Kết quả phát hiện trên ảnh nhiễu; (f) Kết quả phát hiện trên ảnh vào .....	97
Hình 4.1. Tính chất lặp lại của các mẫu chất liệu.....	100

## MỞ ĐẦU

Sự phát triển của máy tính về cả phần cứng và phần mềm đã tạo tiền đề cho nhiều lĩnh vực công nghệ thông tin phát triển, trong đó thị giác máy (computer vision) là lĩnh vực có nhiều nghiên cứu đã và đang được đưa vào ứng dụng một cách hiệu quả. Gần 80% thông tin con người thu nhận được là từ hình ảnh. Nhận dạng tự động, mô tả đối tượng, phân loại phân nhóm các mẫu là những vấn đề quan trọng trong thị giác máy, được ứng dụng trong các ngành khoa học khác nhau. Vấn đề khác, hẹp hơn nhưng được ứng dụng nhiều, đó là việc phát hiện ra đối tượng hoặc một vùng chất liệu trong bức ảnh cho trước [35,37,93,95]. Với sự bùng nổ của thông tin, việc xác định đối tượng hay một mẫu chất liệu nói chung trong ảnh một cách tự động là vấn đề hết sức cần thiết, đặc biệt trong các hệ thống giám sát tự động như hệ thống giám sát vào ra, giám sát giao thông cũng như các hệ thống tự động hóa bao gồm việc xây dựng rô bốt thông minh và trong các hệ thống thực tại ảo [98].

Một cách chung nhất, trong thực tế có thể xem chất liệu được tạo nên từ một hoặc nhiều vật chất. Trong xử lý ảnh, chất liệu của một đối tượng chính là thành phần bao phủ bên ngoài của đối tượng đó, là thành phần không thể thiếu được của mỗi đối tượng. Theo Merriam và Webster [92,93] thì đối tượng là bất kỳ những gì chúng ta cảm nhận được bằng giác quan. Như vậy, chất liệu cũng là bất kỳ nội dung ảnh mà chúng ta cảm nhận được tại các tỉ lệ khác nhau. Theo đó, việc phát hiện mẫu chất liệu cũng bao gồm cả phát hiện vùng ảnh chứa mẫu chất liệu trong một bức ảnh. Nội dung ảnh của một đối tượng chính là nội dung của mẫu chất liệu trong ảnh. Nội dung ảnh có thể được thể hiện bởi các đặc trưng màu sắc, kết cấu (texture), hình dạng và các thông tin không gian. Theo quan điểm này người ta định hướng nghiên cứu và phát triển các hệ thống tra cứu ảnh dựa trên nội dung (Content Based on Image Retrieval – CBIR). Tra cứu ảnh là kỹ thuật tìm và sắp xếp các ảnh theo mức độ tương tự giảm dần với một hoặc nhiều ảnh đầu vào trong một cơ sở dữ liệu ảnh cho trước. Tra cứu ảnh dựa trên nội dung là phương pháp tra cứu dựa vào những thông tin được trích chọn tự động từ ảnh. Nghiên cứu về tra cứu ảnh không những giải quyết những vấn

đề một cách đơn lẻ, mà còn có nhiều nhiệm vụ quan trọng khác như hỗ trợ phát hiện hay nhận dạng đối tượng. Một trong những thách thức đầu tiên là tìm kiếm những ảnh tương tự. Khái niệm tương tự được định nghĩa chủ yếu dựa trên màu sắc, kết cấu và các đặc trưng không gian. Vấn đề này đã được khảo sát sâu vào những năm 90 của thế kỷ trước và đầu những năm 2000, và hiện nay nó vẫn đang được tiếp tục phát triển. Với tra cứu ảnh dựa trên nội dung, Eakins và Graham đã chia thành ba mức, tùy thuộc vào mức độ phức tạp tăng dần đó là: Tìm kiếm ảnh dựa vào các đặc trưng cơ bản như màu sắc, hình dạng, kết cấu và phân bố không gian hay tổ hợp các đặc trưng đó. Các hệ thống tra cứu ảnh thành công ở mức này như: QBIC (Flickner et al 1995), SIMBA (Siggelkow et al 2001), VIPER/GIFT (Muller 2001) hoặc FIRE (Deselaers et al, 2004). Ở mức 1, tất cả các thông tin cần thiết được sử dụng để đánh giá đều được thu thập từ chính bản thân ảnh. Ở mức 2, tìm kiếm đối tượng dựa vào các đặc trưng logic hoặc suy diễn, tra cứu các đối tượng theo một kiểu cho trước, nghĩa là tìm kiếm các thành phần của một loại đối tượng như tìm kiếm các ảnh có “bông hoa”, “con vật” hoặc “da”, “gỗ” v.v hoặc tra cứu các đối tượng riêng hay một đối tượng người cụ thể nào đó, ví dụ tìm các ảnh có ô tô đặc biệt, ở mức này đã có một số công trình nghiên cứu [33,50,58,78,83,97], đặc biệt gần đây nhất là công trình của Alexandra Teynor năm 2009 [92] đã tra cứu một vài loại đối tượng như xe đạp, xe máy. Mức thứ 3 đó là tra cứu dựa trên ngữ nghĩa, tra cứu bằng các sự kiện, đây cũng là một vấn đề khó để giải quyết hoặc các hoạt động như trận bóng đá hay thi Olympic hay tra cứu các bức tranh nói về cảm xúc như tình yêu hay lòng yêu nước v.v. Và hiện nay, bài toán ở mức 3 vẫn chưa có một nghiên cứu nào thậm chí trường hợp thứ hai là chưa thể giải quyết trong tương lai gần.

Tra cứu ảnh dựa trên nội dung chủ yếu quan tâm nghiên cứu ở khía cạnh kỹ thuật, không quan tâm đến khía cạnh người sử dụng. Thay vì người sử dụng phải yêu cầu tìm các bức ảnh có phân phối màu tương tự hay giống với một mẫu chất liệu cho trước, họ sẽ thích hơn khi chỉ cần yêu cầu tìm các bức ảnh có một hoặc một số mẫu chất liệu cho trước nào đó hay tìm kiếm phát hiện một hoặc một vài mẫu chất liệu



nào đó có trong bức ảnh cho trước. Phát hiện mẫu chất liệu trong ảnh hiện đang là một vấn đề cốt lõi trong nhiều hệ thống giám sát tự động [90,98]. Giám sát hỏa hoạn, giám sát giao thông, giám sát theo dõi và bảo vệ vào ra trong các tòa nhà lớn v.v là những bài toán thiết yếu của mỗi quốc gia. Đôi lúc phát hiện được chất liệu thông qua mẫu chất liệu sẽ cho ta phát hiện được đối tượng cần quan tâm, ví dụ nếu phát hiện được chất liệu da mặt thì khả năng trong bức ảnh đó sẽ có mặt người hay có người. Hoặc phát hiện một vùng ảnh nào đó chứa lông Hồ sẽ chỉ ra có đối tượng con Hồ trong ảnh đó v.v. Như vậy, phát hiện mẫu chất liệu trong ảnh không những là một bài toán quan trọng mà còn là một cách tiếp cận mới cho phát hiện đối tượng, là một nghiên cứu hỗ trợ ứng dụng giải quyết bài toán tra cứu ảnh thuộc mức hai theo cách phân chia của Eakins và Graham như đã trình bày ở trên, đây cũng là bước đầu tiên trong các hệ thống nhận dạng. Điều này chứng tỏ được ý nghĩa thực tiễn của bài toán phát hiện mẫu chất liệu trong ảnh.

Mặc khác, ảnh của mẫu chất liệu phụ thuộc nhiều vào các nhân tố: điều kiện chiếu sáng, cấu trúc hình học của bề mặt theo từng tỉ lệ không gian cụ thể cũng như các thuộc tính phản xạ ánh sáng của bề mặt chất liệu phụ thuộc vào hướng chiếu sáng, tỉ lệ và hướng thu nhận v.v. Điều này dẫn đến sự thay đổi lớn trong các thể hiện của mẫu chất liệu, tức là cùng một mẫu chất liệu nhưng thu nhận dưới những điều kiện môi trường khác nhau sẽ trông rất khác nhau (biến thể bên ngoài lớn) hoặc hai mẫu chất liệu khác nhau nhưng trông rất giống nhau (biến thể bên trong nhỏ) [16,32]. Cho đến nay vẫn chưa có một mô hình toán học nào có thể mô tả được các sự thay đổi do môi trường thu nhận ảnh như vậy. Đây chính là vấn đề khó nhất của các nghiên cứu về chất liệu cũng như của bài toán phát hiện mẫu chất liệu trong ảnh mà luận án đặt ra, và hiện nay vẫn còn đang là một thách thức đối với các nhà nghiên cứu [19,20,27,101,102,107]. Hình ảnh dưới đây cho thấy thể hiện mẫu chất liệu thay đổi khi ánh sáng và hướng thu nhận ảnh thay đổi, mỗi chất liệu được thể hiện trên mỗi cột và mỗi hàng thể hiện sự thay đổi theo điều kiện thu nhận.



*Mỗi dòng thể hiện sự thay đổi của mẫu chất liệu trên mỗi cột.*

Nghiên cứu về việc cảm nhận và hiểu được chất liệu đã có từ rất lâu [1,8,88] và chủ yếu cho mục đích xây dựng hay tái tạo lại chất liệu trong đồ họa máy tính. Đối với lĩnh vực thị giác máy, có hai hướng nghiên cứu chính về phát hiện mẫu chất liệu trong ảnh, đó là dựa vào mô hình phản xạ ánh sáng và dựa vào thể hiện ảnh của mẫu chất liệu. Theo hướng thứ nhất, mẫu chất liệu được mô hình hóa bởi các hàm phân phối phản xạ hai chiều BRDF (Bidirectional Reflectance Distribution Function), BTF (Bidirectional Texture Function) và các biến thể của nó [23,24-26,30,80]. BRDF hay BTF chính là ảnh chất liệu được tham số hóa bởi các tham số về ánh sáng và hướng thu nhận của một hàm được xác định trước. Việc nhận dạng hay phát hiện có thể dựa vào các tham số ước lượng được từ các mô hình này trên tập mẫu chất liệu nhưng rất hạn chế về một số điều kiện như ánh sáng, hình học bề mặt và thuộc tính chất liệu [101,102]. Với các mẫu chất liệu trong tự nhiên để sử dụng được mô hình BRDF và BTF đòi hỏi phải huấn luyện tất cả các biến thể của mẫu chất liệu dưới mọi điều kiện ánh sáng và hướng thu nhận ảnh. Việc học như vậy là rất phức tạp, thậm chí không thể thực hiện được bởi vì chúng ta không thể thu thập được một tập các ảnh dưới tất cả các điều kiện ánh sáng với các hướng chụp khác nhau. Tuy nhiên, điều quan trọng nữa là chỉ biết thuộc tính phản xạ của một bề mặt thì vẫn không đủ để quyết định mẫu chất liệu, vì thực tế với một bề mặt trong suốt chúng ta không thể biết nó được làm từ nhựa dẻo, sáp ong hay kính v.v.

Hướng nghiên cứu thứ hai, dựa vào các kiểu đặc trưng ảnh trên cơ sở các đặc trưng địa phương và đặc trưng toàn cục. Với đặc trưng toàn cục, có các phương pháp dựa trên biến đổi miền không gian và tần số. Các phương pháp loại này trích chọn đặc trưng dựa trên biến đổi không gian và tần số như bộ lọc Gabor, bộ lọc Gauss hay phép biến đổi Wavelet, Fourier v.v [104-106]. Một số khác là các phương pháp thống kê, ví dụ điển hình là phương pháp Histogram, Ma trận đồng hiện mức xám (Grey Level Co-occurrence Matrices-GLCM) hay mẫu nhị phân địa phương (Local Binary Patterns-LBP) [77]. Phương pháp GLCM trích chọn các đặc trưng thống kê như độ tương phản, tính đồng đều, độ thô, năng lượng v.v từ đặc tả mối quan hệ không gian của mẫu chất liệu. Các đặc trưng thống kê như ma trận đồng hiện mức xám (GLCM) rất nhạy với sự thay đổi ánh sáng. Tiếp theo là một số phương pháp dựa vào mô hình và xây dựng các tính chất bất biến từ đặc trưng của mô hình, chẳng hạn như mô hình trường Markov ngẫu nhiên (Markov Random Field -MRF) [101], tìm ra những tính chất bất biến ánh sáng. Tuy nhiên, trong trường hợp tổng quát thì các phương pháp thuộc các loại này không bất biến với các phép biến đổi hình học cũng như quang học làm cho việc mô tả và phát hiện mẫu chất liệu không hiệu quả chẳng hạn với các bề mặt có độ lồi lõm lớn thì hướng ánh sáng thay đổi sẽ làm thay đổi thể hiện ảnh của nó rất lớn do bóng (shadow), che khuất (occlusion). Mô hình Fractal sử dụng đặc điểm tự tương tự để tìm ra các đặc trưng bất biến đối với các phép biến đổi affine trong đó quan trọng nhất là bất biến tỉ lệ toàn cục và ánh sáng. Gần đây, các đặc trưng địa phương được quan tâm nghiên cứu nhiều vì thế nó có thể dễ dàng tích hợp các tính chất bất biến vào từng mô tả địa phương tùy thuộc vào từng mục đích của bài toán. Tính địa phương của nó cũng thuận lợi trong việc đối sánh và tìm kiếm mẫu chất liệu trong ảnh. Một số đặc trưng bất biến địa phương tiêu biểu như LBP (2003), LBP-HF (2006), SIFT(1999, 2004) [60,77] v.v, mỗi loại đặc trưng này phù hợp với một số các loại mẫu chất liệu. Các phương pháp loại này sẽ làm giảm bớt số lượng mẫu huấn luyện, thậm chí chỉ cần một mẫu huấn luyện là đủ, nghĩa là thời gian thực hiện nhanh hơn do đó thích hợp cho bài toán phát hiện mẫu chất liệu. Xu hướng nghiên cứu chất liệu, mô hình chất liệu và các phương pháp phát hiện mẫu chất liệu

là các vấn đề nền tảng của thị giác máy và đã được nhiều tác giả nghiên cứu. Hiện tại các nghiên cứu vẫn còn rời rạc chỉ hạn chế trong một số điều kiện cụ thể làm cho việc ứng dụng cũng hạn chế theo. Do đó, việc nghiên cứu và đề xuất các thuật toán, phương pháp phát hiện mẫu chất liệu từ những bức ảnh trong thế giới thực và nhằm mục đích ứng dụng vào lĩnh vực giám sát tự động đòi hỏi các phương pháp giải quyết phải thực hiện thời gian thực. Do tính đa dạng và phức tạp của các loại chất liệu khác nhau trong thế giới thực, vì vậy cho đến hiện nay bài toán này vẫn còn đang được quan tâm nghiên cứu mặc dù hầu hết đều xây dựng các hệ thống nhận dạng chất liệu trong một số các cơ sở dữ liệu chất liệu ứng dụng trong những mục đích cụ thể [27,51,54]. Gần đây nhất, năm 2013 nhóm Edward H. Adelson, Ce Liu, Lavanya Sharan đã đưa ra một phương pháp nghiên cứu nhận dạng loại chất liệu thông dụng dựa vào đặc trưng nhận thức của con người trên cơ sở dữ liệu Flickr Materials Database do họ xây dựng mặc dù cơ sở dữ liệu này không thể hiện được biến thể bên ngoài nhiều mà các bài toán phát hiện cũng như nhận dạng mẫu chất liệu hết sức quan tâm và hơn nữa độ chính xác vẫn còn dưới 50% [89].

Trên cơ sở ý nghĩa thực tiễn và những vấn đề vẫn đang còn nhiều thách thức trong bài toán tra cứu ảnh ở mức hai cũng như việc mô tả và nhận dạng chất liệu của lĩnh vực Thị giác máy, luận án lựa chọn đề tài “***Nghiên cứu phát hiện mẫu chất liệu trong ảnh***”, nhằm nghiên cứu các cách tiếp cận cũng như các phương pháp biểu diễn mẫu chất liệu và tìm kiếm, xác định mẫu chất liệu trong ảnh, đồng thời đề xuất ứng dụng giải quyết bài toán phát hiện ảnh số giả mạo cũng như định hướng nghiên cứu một số bài toán trong lĩnh vực giám sát tự động.

Như đã phân tích ở trên, nghiên cứu về chất liệu thì hầu hết người ta tập trung nghiên cứu theo hai hướng chính đó là mô hình hóa các thuộc tính phản xạ bề mặt và nhận dạng kết cấu 2D và 3D trên cơ sở sử dụng đặc trưng ảnh. Nhưng nhận biết được thuộc tính phản xạ bề mặt vẫn chưa đủ để xác định mẫu chất liệu vì có rất nhiều mẫu chất liệu khác nhau nhưng có cùng thuộc tính phản xạ bề mặt, do đó không thể áp dụng trực tiếp được cho bài toán phát hiện mẫu chất liệu. Nhận biết kết cấu 2D hoặc

3D [24,45,56,82], có nhiều mẫu chất liệu có cùng kết cấu (cùng cấu trúc hình học bề mặt) nhưng được tạo ra từ các chất liệu khác nhau, do đó cũng không thể áp dụng các phương pháp nhận dạng kết cấu cho bài toán phát hiện mẫu chất liệu. Thậm chí người ta đã từng nghiên cứu mối quan hệ giữa đối tượng và chất liệu ví dụ cái bàn có thể làm bằng gỗ, cốc cafe có thể được làm bằng gốm nhưng ngày nay các đối tượng được làm nhân tạo rất nhiều nên quan hệ giữa đối tượng và chất liệu không phải là 1-1, vì vậy theo hướng này vẫn không khả thi cho bài toán của chúng ta. Hơn nữa hầu hết các nghiên cứu về chất liệu đều tập trung vào vấn đề nhận dạng và phụ thuộc vào từng cơ sở dữ liệu cụ thể, trong khi bài toán đặt ra của luận án là tìm kiếm mẫu chất liệu trong ảnh có nghĩa là thông tin về mẫu chất liệu cần tìm và thông tin về ảnh vào vẫn chưa xác định được kích thước, tỉ lệ và các thông số điều kiện thu nhận khác cũng như vị trí của mẫu chất liệu trong ảnh và cho đến hiện nay chưa có một hệ thống nào phát hiện mẫu chất liệu trong ảnh thực sự được đề xuất.

Từ những khó khăn của bản chất bài toán phát hiện cũng như các mô tả mẫu chất liệu như vậy, luận án tập trung nghiên cứu mẫu chất liệu dựa vào thể hiện ảnh của chất liệu để tìm ra các đặc trưng chất liệu bất biến hoặc ít nhạy với một số phép biến đổi hình học và ánh sáng. Đồng thời nghiên cứu các phương pháp phát hiện mẫu chất liệu dựa vào các đặc trưng bất biến đó như bất biến tỉ lệ (gây ra do góc nhìn), bất biến ánh sáng (gây ra do độ chiếu sáng và hướng ánh sáng) dựa vào trích chọn các điểm quan tâm và dựa vào các mô hình [94,99,111,112]. Từ đó, đề xuất một vài kỹ thuật biểu diễn mẫu chất liệu cũng như các thuật toán phát hiện mẫu chất liệu trên mỗi biểu diễn đó, nhằm đánh giá và minh chứng được các đặc trưng này phát hiện được các mẫu chất liệu khá hiệu quả.

Cụ thể, luận án tập trung nghiên cứu các vấn đề sau:

- 1) Tìm các dạng thay đổi ảnh về mặt hình học và quang học đối với các điều kiện thu nhận khác nhau. Nghiên cứu những thuộc tính bề mặt của mẫu chất liệu, từ đó nghiên cứu các kỹ thuật biểu diễn hay mô hình hóa chất liệu thông qua mẫu chất liệu bất biến hoặc ít ảnh hưởng với một số phép biến đổi ảnh.

- 2) Nghiên cứu và đề xuất một số kỹ thuật phát hiện mẫu chất liệu trong ảnh dựa vào các biểu diễn mẫu chất liệu trong những trường hợp có sự thay đổi của môi trường thu nhận ảnh như: nhiễu, ánh sáng, tỉ lệ, bóp méo, che khuất, v.v.
- 3) Nghiên cứu và đề xuất ứng dụng của phát hiện mẫu chất liệu trong ảnh vào một số bài toán hết sức thời sự hiện nay trong giám sát tự động: Phát hiện ảnh số giả mạo.

Với mục đích đặt ra và phân tích các thách thức của bài toán phát hiện mẫu chất liệu trong ảnh, luận án đã đạt được một số kết quả, góp phần nghiên cứu và xây dựng các ứng dụng thế giới thực trong lĩnh vực thị giác máy. Có thể khái quát các kết quả chính của luận án như sau:

- Đề xuất kỹ thuật phát hiện mẫu chất liệu dựa vào đặc trưng bất biến địa phương trên cơ sở phân cụm các mối tương quan hình học của các đặc trưng địa phương để xác định cấu trúc mẫu chất liệu trên ảnh. Kỹ thuật này đã được đăng tải trong Kỷ yếu Hội nghị Quốc Gia về Công nghệ Thông tin và Truyền thông năm 2009, nhằm giải quyết bài toán trong trường hợp có sự thay đổi tỉ lệ, quay và một số phạm vi ánh sáng nhất định. Kết quả thực nghiệm cho thấy kỹ thuật đề xuất khá hiệu quả khi các mẫu chất liệu có kết cấu không mịn, thô, có độ tương phản cao như gỗ, cỏ, vải, v.v. Qua đó, luận án cũng đã đề xuất ứng dụng kỹ thuật này để giải bài toán phát hiện ảnh số giả mạo dạng cắt dán trên cùng một ảnh trong trường hợp vùng giả mạo bị thay đổi hình học mà các kỹ thuật đương thời chưa giải quyết được, kết quả đã được đăng tải trên Tạp chí Tin học và Điều khiển học năm 2010.
- Đề xuất sử dụng đặc trưng nhiễu để biểu diễn và phát hiện mẫu chất liệu. Nhiễu là thành phần không mong muốn, thông thường trong hầu hết các ứng dụng cần thiết phải loại bỏ chúng, song qua nghiên cứu về quá trình thu nhận ảnh thông qua màn phim và giá trị của điểm ảnh được nội suy từ một lân cận do đó chính bản thân mỗi điểm ảnh trên mẫu chất liệu phụ thuộc vào lân cận của nó rất nhiều và có nghĩa nhiễu chất liệu luôn luôn tồn tại. Kỹ thuật này

sử dụng phương pháp học tích lũy các mẫu nhiễu chất liệu và dựa vào phân bố Gauss của độ tương quan để xác định mẫu chất liệu trên ảnh. Thực nghiệm cho thấy khả năng phát hiện các mẫu chất liệu khi có sự thay đổi ánh sáng là rất tốt và kết quả đã được đăng tải ở Tạp chí Khoa học và Công nghệ năm 2010 cùng với Hội nghị FAIR năm 2009.

- Đề xuất một kỹ thuật mô tả và phát hiện dựa vào cấu trúc lặp lại của bản chất chất liệu đó là hình học Fractal. Đề xuất này nhằm giải quyết vấn đề thay đổi tỉ lệ toàn cục của mẫu chất liệu trên ảnh thông qua ý tưởng nén ảnh Fractal. Kỹ thuật được trình bày và đăng tải ở Hội nghị quốc tế ACM (iiWAS-MoMM2011) về Tính toán Thông tin và Truyền thông đa phương tiện năm 2011.

Bố cục của luận án bao gồm phần mở đầu, kết luận và bốn chương nội dung cùng với phụ lục và tài liệu tham khảo.

Chương 1: Trình bày một cái nhìn thống nhất về bài toán phát hiện mẫu chất liệu đồng thời trình bày một số giải pháp mô tả và phát hiện mẫu chất liệu, cũng như các cách tiếp cận để biểu diễn và phát hiện mẫu chất liệu trong ảnh.

Chương 2: Trình bày về đặc trưng bất biến địa phương và đề xuất sử dụng đặc trưng bất biến địa phương cho việc tìm trực tiếp mẫu chất liệu xuất hiện trong ảnh dựa vào cấu trúc tương quan hình học của các đặc trưng trong mẫu chất liệu. Trong chương này cũng đề xuất ứng dụng phát hiện mẫu chất liệu cho bài toán phát hiện ảnh số giả mạo dạng cắt dán trong trường hợp vùng nghi ngờ giả mạo bị thay đổi tỉ lệ và quay. Các đề xuất được cài đặt và đánh giá thử nghiệm cho thấy kết quả phát hiện khá hiệu quả.

Chương 3: Luận án trình bày về khái niệm nhiễu, sự hình thành của nhiễu trong quá trình thu nhận ảnh và chứng minh đặc trưng ảnh của mẫu chất liệu phụ thuộc nhiễu vào đặc trưng nhiễu của nó. Từ đó đề xuất một kỹ thuật mới cho phát hiện mẫu chất liệu dựa trên đặc trưng nhiễu này. Kỹ thuật được cài đặt và kết quả

thực nghiệm cho thấy kỹ thuật đề xuất phát hiện mẫu chất liệu khá tốt trong trường hợp ảnh bị thay đổi bởi ánh sáng và mẫu chất liệu có cấu trúc hình học bề mặt thấp.

Chương 4: Trình bày biểu diễn mẫu chất liệu dựa vào hình học Fractal trên cơ sở các hệ hàm lặp IFS mô tả tính chất tự tương tự tại các tỉ lệ khác nhau trên các vùng khác nhau của mẫu chất liệu nhằm xây dựng mô hình Fractal cho mẫu chất liệu. Đồng thời đề xuất hai thuật toán biểu diễn và phát hiện mẫu chất liệu dựa trên tiếp cận Fractal.



# **Chương 1. TỔNG QUAN VỀ PHÁT HIỆN MẪU CHẤT LIỆU TRONG ẢNH**

Phần này trình bày khái quát về chất liệu, mẫu chất liệu và bài toán phát hiện mẫu chất liệu trong ảnh, các cách tiếp cận cho bài toán đặt ra và các nghiên cứu liên quan nhằm mục đích làm cơ sở cho các nghiên cứu chính trong các chương sau.

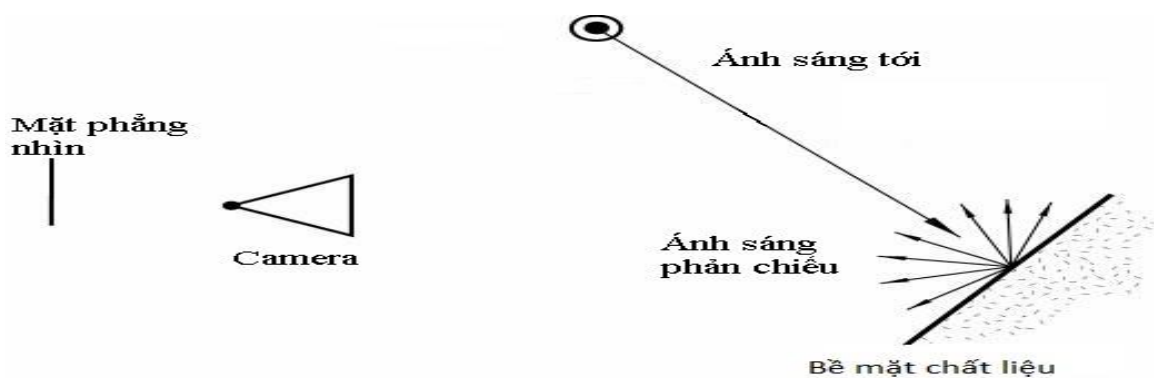
## **1.1. Chất liệu và bài toán phát hiện mẫu chất liệu trong ảnh**

### **1.1.1. Chất liệu và mẫu chất liệu trong ảnh**

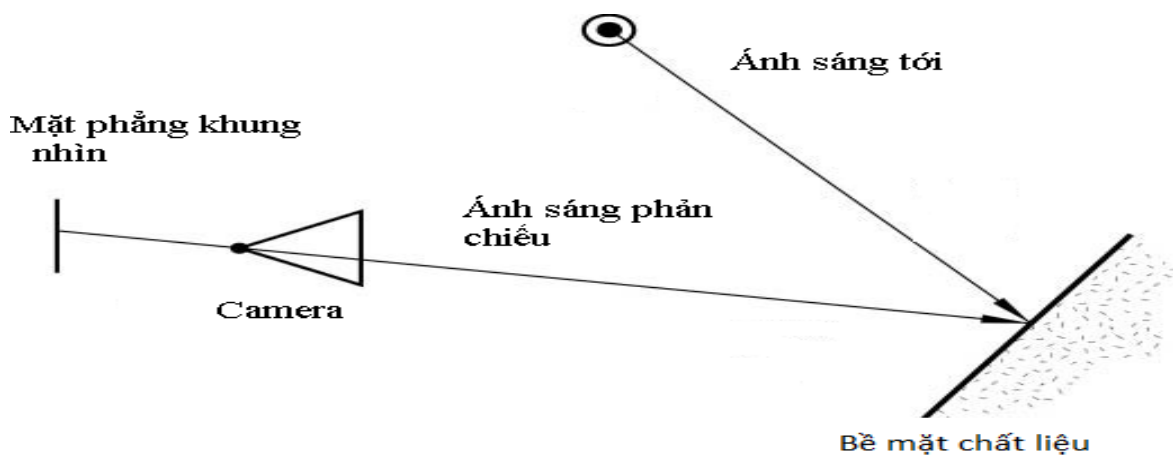
Chất liệu là một dạng tồn tại vật chất nào đó và được cấu tạo từ một hoặc nhiều thành phần hóa học khác nhau theo một nguyên tắc nhất định, chúng có cùng tính chất vật lý, ví dụ gỗ được cấu tạo chủ yếu từ xenlulo và một số chất hóa học khác. Tùy thuộc vào sự sắp xếp các thành phần hóa học để cấu tạo nên những chất liệu khác nhau, chẳng hạn có nhiều chất liệu gỗ khác nhau về sự lặp lại của các đường vân, màu sắc, độ cứng, chịu nhiệt, hấp thụ và phản xạ ánh sáng khác nhau, nên không thể có một định lượng vật lý cụ thể cho một chất liệu mà chỉ là một khoảng xác định. Do đó, chất liệu có thể được hiểu dựa trên các phương diện khác nhau theo từng lĩnh vực khoa học nghiên cứu.

Trong xử lý ảnh và thị giác máy, chất liệu được thể hiện dưới dạng ảnh, nghĩa là chất liệu chỉ có thể cảm nhận và nhận biết được chỉ thông qua quan sát, phân tích sự tương tác của chúng đối với ánh sáng. Mỗi chất liệu khi quan sát được đó là do ánh sáng mặt trời chiếu vào và phản xạ đến mắt người, tùy thuộc từng loại chất liệu mà sự phản xạ ánh sáng sẽ khác nhau. Chẳng hạn có chất liệu thì phản xạ ánh sáng mạnh như gương soi hoặc mặt phẳng nhẵn, những chất liệu khác thì phản xạ ánh sáng yếu hoặc thậm chí còn gần như là trong suốt, có những chất liệu thì phản chiếu ánh sáng theo một hướng những chất liệu thì phản chiếu ánh sáng theo mọi hướng hoặc theo nhiều hướng khác nhau. Màu mà mắt người quan sát được chính là màu ánh sáng phản xạ từ chất liệu, nên ảnh chất liệu nhận được dựa vào lượng ánh sáng phát tán từ chất liệu đến mắt người và có sự ghi nhận khác nhau. Trong thực tại ảo, người ta đã

tiếp cận dựa vào các tính chất đó của chất liệu trong thực tế, vận dụng trong kỹ thuật che phủ và camera cũng giống như mắt người quan sát vậy. Ánh sáng phản chiếu từ chất liệu được cụ thể thành từng loại ánh sáng riêng biệt như ánh sáng phản chiếu khúc xạ, ánh sáng phản chiếu toàn bộ, ánh sáng phản chiếu từ môi trường xung quanh v.v. Chính vì những yếu tố của thực tế mà trong môi trường thực tại ảo các loại ánh sáng đó được cụ thể bằng các loại màu phản chiếu từ chất liệu. Khái niệm chất liệu dùng trong kỹ thuật che phủ là sự tổng hợp của ánh sáng và chất liệu cấu tạo để tạo nên bề mặt chất liệu và cả hai nhân tố trên đều là rất quan trọng trong việc quyết định các bề mặt chất liệu sẽ được xuất hiện như thế nào, nghĩa là từ sự xuất hiện của bề mặt chất liệu cho trước ta có thể xác định chất liệu cấu tạo nên đối tượng đó thông qua nguyên tắc tổng hợp và phân tích ánh sáng nhất định. Ánh sáng gồm cả thành phần trong môi trường tác động lên chất liệu cấu tạo để nhận được ảnh chất liệu. Thực chất, chất liệu thu nhận được thông qua quá trình thu nhận ảnh và chất liệu cấu tạo sai khác nhau do môi trường ánh sáng và quá trình thu nhận. Thể hiện ảnh của chất liệu phụ thuộc vào bản chất bề mặt của chất liệu cấu tạo trong đó có hình dạng kết cấu cụ thể, thuộc tính ánh sáng và thuộc tính hình học của môi trường thu nhận. Hình 1.1 và hình 1.2 bên dưới minh họa quá trình thu nhận chất liệu phụ thuộc sự phản xạ ánh sáng.



Hình 1.1. Sự phản chiếu theo một hướng (Specular Reflection)



Hình 1.2. Sự phản chiếu khuếch tán (Diffuse Reflection)

Nếu chúng ta biết được các thuộc tính cụ thể của các thành phần bên ngoài đó và bên trong đó, thì việc phát hiện chất liệu của đối tượng hay chất liệu cấu tạo dễ dàng hơn rất nhiều.

Ngày nay, chất liệu trong thực tế rất đa dạng và phong phú gồm chất liệu tự nhiên và chất liệu nhân tạo, việc nghiên cứu bài toán phát hiện mẫu chất liệu trong xử lý ảnh còn phụ thuộc vào điều kiện thu nhận ảnh do đó càng phức tạp hơn. Dù chất liệu được xét trên phương diện nào thì nó cũng phản ánh được phần nào đó của chất liệu thực. Con người có nhiều cách để nhận biết được chất liệu theo nhiều phương diện khác nhau và trong thị giác máy cũng vậy, để thuận tiện cho việc nghiên cứu các phương pháp biểu diễn và phát hiện chất liệu thông qua mẫu chất liệu, mẫu chất liệu có thể chia làm hai loại:

- *Mẫu chất liệu có mô hình biểu diễn*: Là các mẫu chất liệu được biểu diễn bằng các quy tắc hoặc các công thức do chuyên gia nghiên cứu về chất liệu đó đưa ra.
- *Mẫu chất liệu không có mô hình biểu diễn*: Trong thực tế có rất nhiều mẫu chất liệu không có mô hình biểu diễn. Trường hợp này mẫu chất liệu sẽ được biểu diễn dưới dạng một hoặc một tập các ảnh của chất liệu đó. Mẫu chất liệu có tính vùng, có sự lặp lại, có sự tương đồng về các thuộc tính hình ảnh. Hình 1.3 là một số mẫu chất liệu không có mô hình biểu diễn.



Hình 1.3. Một số mẫu chất liệu thông thường

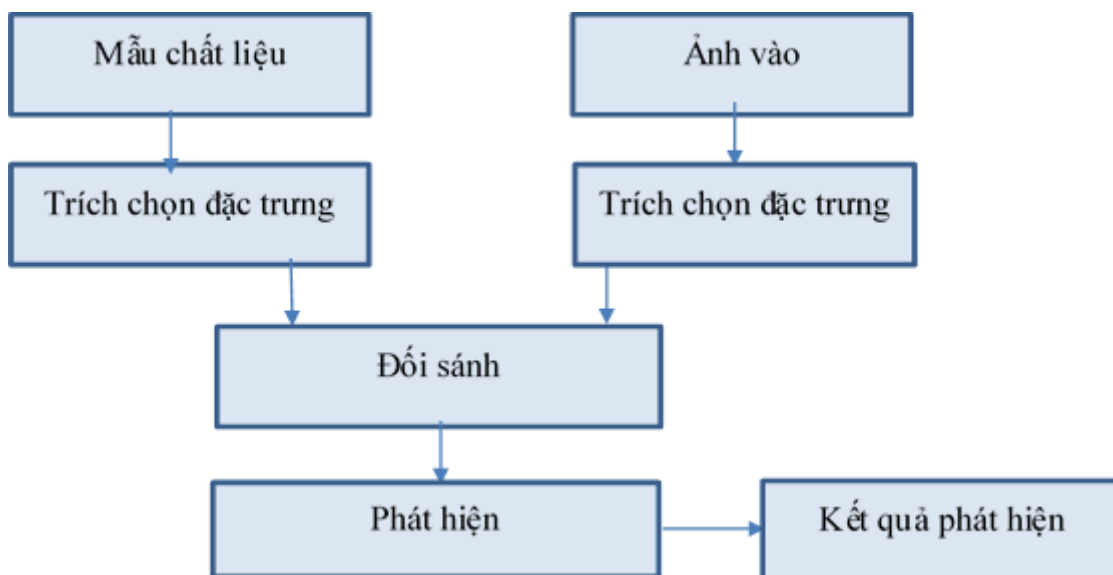
### 1.1.2. Bài toán phát hiện mẫu chất liệu trong ảnh

Đối với con người, để phát hiện chất liệu trong ảnh chúng ta có thể nhìn và so sánh chúng với chất liệu đã được chúng ta hình dung và mô tả theo nhiều hướng khác nhau hoặc sờ mó vào chúng. Nhưng trong xử lý ảnh và thị giác máy, hay nói khác hơn là trên phương diện quan sát và nhận biết, chất liệu được thể hiện dưới dạng mẫu chất liệu, nó là ảnh của chất liệu thật, nó phản ánh một số thuộc tính của chất liệu thực tế. Như vậy bài toán phát hiện chất liệu trong ngữ cảnh xử lý ảnh và thị giác máy sẽ chính là bài toán phát hiện mẫu chất liệu trong ảnh, tức là tìm xem trong một bức ảnh cho trước có mẫu chất liệu nào đó hay không. Phát biểu bài toán cụ thể như sau:

Cho trước một bức ảnh  $\mathcal{J}$  và một mẫu chất liệu  $\mathcal{M}$ . Hãy tìm kiếm và xác định các vùng chứa mẫu chất liệu  $\mathcal{M}$  trên ảnh  $\mathcal{J}$ ?

Phát hiện mẫu chất liệu hay nói khác hơn là tìm và xác định vị trí và kích thước của một mẫu chất liệu cho trước trong một bức ảnh nào đó. Trong trường hợp mẫu chất liệu có mô hình biểu diễn, phát hiện mẫu chất liệu được thực hiện đơn giản bằng cách áp dụng công thức của mẫu chất liệu  $\mathcal{M}$  lên ảnh chúng ta sẽ nhận được các vùng ảnh chứa mẫu chất liệu đó. Đối với loại mẫu chất liệu không có mô hình biểu diễn, việc phát hiện mẫu chất liệu chính là tìm ảnh con trong ảnh lớn. Việc tìm ảnh con trong ảnh lớn đối với con người thì hết sức dễ dàng, nhưng đối với máy tính thì vẫn đang là vấn đề hết sức khó khăn vì thông thường các ảnh con đều có sự thay đổi quang học và hình học gây ra do nhiều yếu tố như: thuộc tính phản xạ bề mặt, ánh sáng, hướng thu nhận và cấu trúc bề mặt v.v làm cho các ảnh con có thể có tỉ lệ, hướng (bi quay), màu sắc khác nhau và khác với ảnh mẫu chất liệu.

### Sơ đồ hệ thống phát hiện mẫu chất liệu trong ảnh:



Hình 1.4. Sơ đồ hệ thống phát hiện mẫu chất liệu tổng quát

#### 1.1.3. Các thách thức của phát hiện mẫu chất liệu trong ảnh

Khó khăn lớn nhất trong bài toán phát hiện mẫu chất liệu trong ảnh đó là việc mô hình hóa hay biểu diễn mẫu chất liệu sao cho có thể phát hiện được vùng chất liệu trong vô vàn các thể hiện khác nhau do bản chất bề mặt cấu tạo, thuộc tính phản xạ của bề mặt, ánh sáng và hướng thu nhận ảnh. Thể hiện của chất liệu phụ thuộc vào bề mặt chất liệu cấu tạo, ánh sáng và hướng thu nhận ảnh. Mẫu chất liệu chỉ thể hiện trên mặt phẳng ảnh 2D, trong khi chất liệu thực tế lại trong không gian 3D, do đó mẫu chất liệu không thể phản ánh được toàn bộ các thuộc tính của chất liệu thực tế.

- **Thuộc tính phản xạ bề mặt của chất liệu:** Bề mặt chất liệu cấu tạo phản ánh thuộc tính phản xạ của bề mặt chất liệu, nhưng thực tế thuộc tính phản xạ của bề mặt không đủ để xác định chất liệu ví dụ một bề mặt trong suốt có thể là kính, nhựa dẻo hay sáp ong v.v.
- **Ánh sáng, hướng thu nhận và cấu trúc hình học của bề mặt:** Thể hiện của chất liệu phụ thuộc nhiều vào điều kiện thu nhận. Hình 1.5 minh họa thể hiện

khác nhau của cùng một chất liệu. Hướng thu nhận ảnh tạo ra các sự thay đổi hình học như: bị quay, bị thay đổi tỉ lệ, bị bóng do thay đổi hướng ánh sáng v.v.



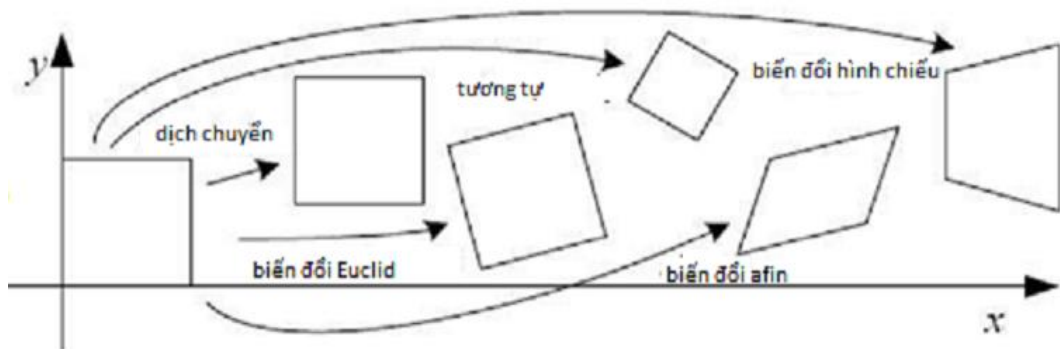
*Hình 1.5. Các thể hiện khác nhau của cùng một mẫu chất liệu*

- **Thiếu mô hình toán học:** Khi mô hình hóa một vấn đề chúng ta luôn phải tạo các giả thiết cho một số điều kiện để đảm bảo vấn đề tính toán là khả thi. Nhưng trong thực tế, rất nhiều điều kiện như phi tuyến, không phẳng, hoặc các phụ thuộc thống kê không giả thiết được.

Do đó, để giải quyết khó khăn này người ta phải xây dựng được mô hình 3D [82,84], hoặc cần phải lấy đầy đủ tất cả các mô hình 2D đặt vào một quan hệ nào đó như Thomas Leung và các cộng sự năm 2003. Đơn giản hơn có thể giả sử các mẫu chất liệu phải phẳng hoặc ít nhất các phần của mẫu chất liệu phải phẳng. Nếu chỉ xét các vùng nhỏ trên bề mặt mẫu chất liệu là trơn thì giả thiết này gần xấp xỉ đúng. Các phần của mẫu chất liệu có thể chịu một số phép biến đổi giữa các ảnh. Để đơn giản giả sử các camera là vô hạn (phép chiếu song song từ camera đến đối tượng) khi đó các phép biến đổi chỉ là các phép biến đổi afin và như thế chúng ta có thể thực hiện mô tả được bằng toán học. Tiếp theo có thể biểu diễn mối quan hệ các tọa độ của một điểm  $x = (x, y)^T \in \mathbb{R}^2$  trong ảnh và được biến đổi thành  $x'$  trong ảnh khác bởi  $x' = Tx + t$ . Trong đó  $t \in \mathbb{R}^2$  là phép dịch chuyển. Ma trận  $T \in \mathbb{R}^{2 \times 2}$  miêu tả các kiểu biến đổi:

- $T = 1_2$       Phép dịch chuyển
- $T^T T = 1_2$     Phép biến đổi Euclid
- $T^T T = \kappa 1_2$     Phép biến đổi tương tự
- $\det(T) \neq 0$     Phép biến đổi afin

$1_n$  kí hiệu ma trận đơn vị  $n$  chiều. Sự thay đổi hình học làm cho việc so sánh các ảnh thậm chí giống nhau cũng trở nên rất khó vì các tham số phép biến đổi không được biết.



Hình 1.6. Các phép biến đổi hình học trong quá trình thu nhận ảnh

- **Tỉ lệ giữa mẫu chất liệu và ảnh vào không tương thích:** Mẫu chất liệu xuất hiện trong ảnh vào quá nhỏ trong khi ảnh vào chứa nhiều mẫu chất liệu khác nổi trội hơn. Điều này làm cho việc phát hiện rất khó khăn vì chúng ta không biết được tỉ lệ thu nhận của mẫu chất liệu và ảnh vào.

## 1.2. Các cách tiếp cận phát hiện mẫu chất liệu trong ảnh

Phát hiện mẫu chất liệu trong ảnh bao gồm hai pha chính: pha thứ nhất là mô hình hóa hay biểu diễn mẫu chất liệu bằng các đặc trưng phải có tính phân biệt cao, đặt biệt là giữa các mẫu chất liệu khác nhau nhưng cũng phải đảm bảo sự giống nhau của các thể hiện khác nhau cho cùng mẫu chất liệu và pha thứ hai là phát hiện dựa vào biểu diễn đó của mẫu chất liệu trong ảnh. Thường hai pha này có mối quan hệ tương quan lẫn nhau, nếu việc mô hình hóa hay biểu diễn cho mẫu chất liệu tốt thì việc phát hiện đơn giản hơn và ngược lại. Việc mô hình hóa hay biểu diễn mẫu chất

liệu có rất nhiều nghiên cứu dựa vào các thuộc tính chất liệu, đặc biệt là thuộc tính kết cấu bề mặt, cũng theo đó nghiên cứu và tìm kiếm các tính chất bất biến hoặc ít nhạy đối với sự thay đổi của môi trường thu nhận cụ thể là ánh sáng và hướng thu nhận. Có rất nhiều chất liệu trong cuộc sống thực tế với các đặc tính bề mặt cụ thể khác nhau, ví dụ vỏ cây thì sần sùi không phẳng; lông, vải hay nước thì dễ biến dạng; gỗ, đá có bề mặt không đồng nhất; kính phụ thuộc vào hướng thu nhận ánh sáng rất lớn v.v. Do đó, các cách tiếp cận để phát hiện chất liệu cũng phụ thuộc vào từng kiểu bề mặt mẫu chất liệu khác nhau. Trên cơ sở đề xuất các phương pháp biểu diễn hay phát hiện ảnh nói chung và mẫu chất liệu nói riêng, cần phải xây dựng các đặc trưng bất biến hay nghiên cứu các hàm phân loại có tính chất bất biến. Dựa trên các nghiên cứu về xử lý ảnh số và đặc biệt là các nghiên cứu về kết cấu bề mặt, hướng ánh sáng cũng như cơ sở toán học của bất biến hình học và quang học, có thể phân tích các phương pháp phát hiện mẫu chất liệu dựa vào các cách tiếp cận dưới đây.

### **1.2.1. Tiếp cận dựa vào đặc trưng địa phương**

Tiếp cận này dựa trên cơ sở lựa chọn các điểm nổi bật và xây dựng các mô tả bất biến đối với một số phép biến đổi ảnh nào đó cho các điểm nổi bật này để tạo ra các đặc trưng bất biến địa phương thông qua các đặc trưng mức thấp của nội dung ảnh như: cường độ sáng, màu sắc, biên, cạnh hoặc các đáp ứng lọc của các phép biến đổi miền không gian tần số. Các đặc trưng này có tính phân biệt cao và sử dụng chúng trong việc đối sánh, nhận dạng hay tra cứu thông qua nội dung ảnh. Đây là một cách tiếp cận đầy hứa hẹn để giải quyết nhiều bài toán Thị giác máy những năm gần đây trong nhận dạng đối tượng [7,10], nhận dạng kết cấu bề mặt [63,64], tra cứu ảnh [83,85,86,100], nhận dạng loại đối tượng [29,36,37], nhận dạng chất liệu [27,51] v.v. Tiếp cận dựa vào đặc trưng địa phương để phát hiện mẫu chất liệu trên cơ sở tìm kiếm và xác định các điểm nổi bật cho mẫu chất liệu và xây dựng các mô tả bất biến với nhiều nhất các thay đổi có thể cho mỗi điểm nổi bật đã xác định được để tạo thành các đặc trưng địa phương và phát hiện mẫu chất liệu bằng cách đối sánh trực tiếp mỗi đặc trưng địa phương trên mẫu chất liệu với các đặc trưng trên ảnh. Tiếp cận này phù



hợp với các mẫu chất liệu có thuộc tính kết cấu bề mặt cao [54]. Tùy thuộc vào cách chọn điểm nổi bật và cách xây dựng mô tả bất biến đối với các phép biến đổi nào sẽ có các loại đặc trưng địa phương khác nhau. Với cách tiếp cận địa phương, cũng có nhiều nghiên cứu mô tả mẫu chất liệu bằng các thành phần cơ bản và mối quan hệ sắp xếp không gian giữa chúng. Mỗi chất liệu có thuộc tính cấu trúc nhất định, do đó có thể có nhiều phương pháp phát hiện cấu trúc chất liệu. Thuộc tính cấu trúc của các phần tử chất liệu đã được sử dụng rất thành công. Việc xây dựng các cấu trúc của từng loại phần tử chất liệu có thể được mô hình hóa bằng phương pháp hình học, phương pháp học máy như SVM [9]. Phương pháp có cấu trúc phát hiện tốt cho các trường hợp chất liệu không phẳng, có tính gồ ghề như vỏ cây, gỗ, đá.v. v. Vấn đề chính của loại phương pháp này là phải tìm kiếm các phần tử chất liệu phù hợp. Sau đây là một số các phương pháp trích chọn đặc trưng địa phương.

#### **1.2.1.1. Phương pháp dựa trên độ cong của đường biên**

Đường biên là sự liên kết tập tất cả các điểm biên liên tiếp. Các điểm giao, điểm cắt là nơi có sự thay đổi lớn của tín hiệu ảnh hay còn gọi là các điểm nổi bật. Do đó việc trích chọn các điểm dọc theo đường biên với độ cong lớn là một trong số các chiến lược hiệu quả trong việc lựa chọn và xây dựng các đặc trưng địa phương. Độ cong của một đường biên thường được xác định bằng tỉ số giữa vector tiếp tuyến đơn vị với độ dài của đường cong. Các nghiên cứu điển hình thuộc loại này như [3] và một số cải thiện để giải quyết vấn đề thay đổi tỉ lệ luôn luôn tồn tại trong các biểu diễn ảnh [61].

#### **1.2.1.2. Phương pháp dựa trên cường độ ảnh**

Các phương pháp thuộc loại này dựa trên đạo hàm bậc nhất và đạo hàm bậc hai ảnh hoặc sử dụng các phương pháp heuristic để tìm các điểm có sự thay đổi cường độ lớn. Các phương pháp thuộc loại này cũng dựa trên hai tiếp cận đó là tiếp cận sai phân và tiếp cận Gradient. Tiếp cận sai phân điển hình là phương pháp Hessian bất biến với phép quay ảnh được đề xuất bởi Beaudet [6]. Nghiên cứu này dựa trên khai triển Taylor bậc hai cụ thể là ma trận Hessian, dựa trên định thức của ma trận Hessian

để tính các điểm cực đại, các điểm cực đại này chính là các cấu trúc ảnh cần trích chọn. Tiếp cận dựa trên Gradient, điển hình là phương pháp Hariss do Hariss và Stephent đề xuất [42]. Sau đó cũng được nhiều nghiên cứu cải thiện dựa trên tiếp cận này [65,69] như phương pháp Hassian - Laplace, Hariss - Laplace. Một số khác kết hợp tìm kiếm các điểm cực trị trên không gian tỉ lệ với các mô tả không gian như lược đồ gradient của các vùng xung quanh một tập các điểm lân cận như phương pháp SIFT của David Lowe [60].

### 1.2.1.3. Phương pháp định hướng bất biến với các phép biến đổi

Một số phương pháp thuộc loại này như: Phương pháp đa tỉ lệ, phương pháp bất biến tỉ lệ, phương pháp bất biến afin v.v. Với cách tiếp cận này, cũng có nhiều nghiên cứu có thể mô tả mẫu chất liệu bằng các thành phần cơ bản và mối quan hệ sắp xếp không gian giữa chúng. Mỗi mẫu chất liệu có thuộc tính cấu trúc nhất định, do đó có thể nhiều phương pháp phát hiện cấu trúc cho mẫu chất liệu. Thuộc tính cấu trúc của các phần tử chất liệu đã được sử dụng rất thành công trong phân lớp mẫu chất liệu [107-109]. Việc xây dựng các cấu trúc của từng loại phần tử chất liệu có thể được mô hình hóa bằng phương pháp cấu trúc hình học, phương pháp học máy như SVM. Phương pháp có cấu trúc phát hiện tốt cho các trường hợp chất liệu không phẳng, có tính gồ ghề như vỏ cây, gỗ, đá v. v. Vấn đề chính của loại phương pháp này là phải tìm kiếm các phần tử chất liệu phù hợp.

### 1.2.1.4. Phương pháp tỉ lệ chu vi và diện tích

Sử dụng chu vi, diện tích để làm các phần tử chất liệu [38,39,57], chu vi bất biến với phép quay, phép dịch chuyển và một cấu trúc trên các phần tử chất liệu này có thể được xác định bằng:

$$r = \frac{P^2}{4\pi S} \quad (1.1)$$

Trong đó  $P$  là chu vi và  $S$  là diện tích. Đặc trưng cấu trúc này không những bất biến với phép quay, dịch chuyển mà còn bất biến với tỉ lệ.

### **1.2.1.5. Phương pháp cấu trúc hình học**

Mô tả cấu trúc chất liệu bằng cấu trúc tô pô, phép biến đổi Hough là một trong các kỹ thuật tìm ra cấu trúc chất liệu dựa trên việc xấp xỉ phương trình đường thẳng, hay một đường cong bất kỳ. Trong thời gian gần đây, việc tìm các phân tử chất liệu có thể dựa vào mô tả địa phương là một trong những cách tiếp cận hứa hẹn nhất. Cách tiếp cận này dựa trên cơ sở tìm kiếm các điểm đặc biệt nổi bật chứa nhiều thông tin nhất và thông tin ở đây chính là yêu cầu cần thiết cho từng ứng dụng của bài toán, sau đó xây dựng các mô tả cho chất liệu từ các điểm nổi bật đã tìm được. Ưu điểm của cách tiếp cận này không những làm giảm thiểu không gian tính toán mà còn giải quyết được vấn đề thách thức trong phát hiện chất liệu như sự thay đổi hình học và quang học. Một số phương pháp thuộc loại này thường được gọi chung là phương pháp đặc trưng địa phương như phương pháp biến đổi đặc trưng bất biến tỉ lệ (Scale Invariant Feature Transform) [60], Harris, Hessian, Harris/ Hessian – Laplace/ Afin.

### **1.2.2. Tiếp cận dựa vào đặc trưng toàn cục**

Mẫu chất liệu có tính vùng, tính thống kê toàn cục, do đó có thể nghiên cứu các phương pháp phát hiện mẫu chất liệu dựa trên các đặc trưng toàn cục. Các đặc trưng toàn cục chủ yếu được xây dựng bằng công cụ thống kê mô tả mẫu chất liệu một cách gián tiếp như đặc trưng phân phối thống kê bậc một và bậc hai mối quan hệ giữa các mức xám của mẫu chất liệu. Thống kê bậc một phân tích trên từng pixel trên trung bình, phương sai. Các thống kê bậc hai là các phân tích trên các lân cận của pixel. Sau đó trích chọn những đặc trưng có tính chất bất biến trên cơ sở các thống kê đã tìm được. Có rất nhiều phương pháp thuộc tiếp cận này như ma trận đồng hiện, mẫu nhị phân cục bộ (Local Binary Patterns –LBP), hiệp phương sai, ma trận tương quan v.v. Ma trận đồng hiện là thống kê bậc hai điển hình, các ma trận này thống kê tần suất xuất hiện của 2 mức xám lân cận theo hướng và khoảng cách. Từ các ma trận tìm được, chúng ta có thể rút ra các đặc trưng như tính thô, độ mịn, độ tương phản v.v. Phương pháp ma trận đồng hiện là một trong các phương pháp thống kê hữu hiệu và được quan tâm nhiều trong thời gian gần đây. Tiếp cận này rất phù hợp

để phát hiện các loại chất liệu có tính ngẫu nhiên không đồng đều như cỏ, thảm v.v. Một số phương pháp thuộc cách tiếp cận này được trình bày dưới đây.

### 1.2.2.1. Phương pháp lược đồ màu

Trong nhiều ứng dụng, lược đồ màu được chứng minh là cách tiếp cận hữu dụng và nhanh nhất. Để tính toán lược đồ màu chỉ cần xét trên các pixel đơn mà không quan đến mối quan hệ giữa chúng, vì vậy lược đồ màu chỉ phản ánh thông tin thống kê bậc nhất của chất liệu. Lược đồ màu bất biến với phép dịch chuyển, phép quay. Các đặc trưng thường được trích chọn trên lược đồ màu là trung bình, phương sai v.v. Có thể sử dụng các phương pháp so lược đồ màu để nhận được các đặc trưng như chuẩn L1, chuẩn L2, giao hiệu các lược đồ màu v.v. Đặc trưng này được dùng nhiều trong các hệ thống tra cứu ảnh và phân đoạn ảnh y tế.

### 1.2.2.2. Phương pháp ma trận đồng hiện mức xám

Ngày nay, ma trận đồng hiện là một trong những công cụ được nghiên cứu và sử dụng nhiều trong phân tích texture được đưa ra đầu tiên bởi Harlick và các cộng sự. Nhóm nghiên cứu này minh chứng khả năng ứng dụng của nó vào việc phân loại ảnh đá, kính và phân đoạn ảnh vệ tinh [54,81]. Các ma trận đồng hiện dựa trên các thống kê bậc hai nên đặc tả được mối quan hệ thông gian giữa hai mức xám bằng vectơ thay thế, có thể trích chọn một số đặc trưng từ các ma trận đồng hiện như năng lượng, entropy, độ tương phản, tính đồng đều v.v.

### 1.2.2.3. Phương pháp mẫu nhị phân địa phương

LBP [77] mô tả quan hệ của một lân cận 8 pixel cho một điểm ảnh, nếu giá trị mức xám của pixel láng giềng cao hơn hoặc bằng giá trị mức xám của pixel trung tâm thì nó đặt bằng 1, ngược lại được đặt bằng 0, kết quả ta thu được một số nhị phân.

Cụ thể:

$$LBP_{R,N}(x, y) = \sum_{i=0}^{N-1} s(n_i - n_c) * 2^i \quad \text{trong đó } s(a) = \begin{cases} 1 & \text{nếu } a \geq 0 \\ 0 & \text{nếu } a < 0 \end{cases} \quad (1.2)$$

$n_c$  là mức xám pixel trung tâm,  $n_i$  là các mức xám của các pixel trong lân cận  $N$  láng giềng với bán kính  $R$ . Phương pháp mẫu nhị phân địa phương (LBP) là một cách phương pháp bất biến quay và ít nhạy với ánh sáng. Phương pháp LBP đã được ứng dụng trong các nghiên cứu của một số loại mẫu chất liệu có cấu trúc bề mặt khá tương đồng như da mặt.

#### 1.2.2.4. Phương pháp dựa vào nhiễu chất liệu

Về mặt vật lý, mỗi chất liệu sẽ có các thuộc tính khác nhau và chắc chắn sẽ có những phương pháp cảm nhận hay phát hiện khác nhau. Song chúng ta chỉ xét trên dữ liệu ảnh, mà chất liệu trong ảnh và chất liệu thực tế là không hoàn toàn giống nhau, tùy thuộc vào tác động ảnh hưởng của môi trường thu nhận ảnh. Các yếu tố môi trường này bao gồm: độ chiếu sáng, các phép biến đổi hình học như phép thay đổi tỉ lệ, phép quay, phép biến đổi affine và các dạng méo mó khác, đồng thời cũng do các loại nhiễu ngẫu nhiên, nhiễu thiết bị điện tử gây ra.

Cũng theo dòng nghiên cứu đó, chúng tôi nhận thấy rằng với chất liệu có cấu trúc tự nhiên đặc trưng nhiễu được xuất hiện rõ nhất. Nhiễu cũng là một trong những vấn đề được nhiều quan tâm, bởi vì chính nó đã làm ảnh hưởng đến hầu hết các thiết kế và xây dựng các hệ thống thông minh, đặc biệt là trong xử lý ảnh và thị giác máy với dữ liệu xử lý hoàn toàn là những bức ảnh. Ảnh không tránh khỏi được nhiễu, nên việc giải quyết các bài toán trên dữ liệu ảnh luôn gặp khó khăn. Trong một số trường hợp thì nhiễu lại là yếu tố có lợi trong việc phân tích và phát hiện chất liệu. Như chúng ta đã biết, nhiễu trong một bức ảnh là do tác động của nhiều thành phần như môi trường, thiết bị thu nhận và đặc biệt là nhiễu do chính bản thân đối tượng tạo ra. Mỗi loại chất liệu khác nhau sẽ có lượng nhiễu sinh ra khác nhau ví dụ như trong hình 1.7 ở dòng trên là các mẫu chất liệu và dòng bên dưới là mẫu nhiễu của chất liệu tương ứng. Luận án đề xuất một kỹ thuật phát hiện dựa trên việc trích chọn và phân tích đặc tính nhiễu của các thể hiện ảnh khác nhau cho cùng một mẫu chất liệu, sau đó tạo mô tả đặc trưng nhiễu chất liệu cho mẫu chất liệu đó. Tìm kiếm mẫu chất liệu

dựa vào đặc trưng nhiễu có thể làm giảm thiểu sự ảnh hưởng của điều kiện ánh sáng, đây là một trong những vấn đề thách thức cho bài toán phát hiện mẫu chất liệu.



Hình 1.7. Nhiễu mẫu chất liệu, dòng trên: các ảnh gốc;  
dòng bên dưới: các ảnh nhiễu tương ứng

Nói chung, nhiễu được xem như thành phần không mong muốn có trong ảnh. Nhiễu là một hiện tượng ngẫu nhiên luôn luôn có mặt trên mọi hệ thống xử lý tín hiệu thực. Nhiễu xuất hiện trong ảnh bởi nhiều nguyên nhân như do sự thay đổi độ nhạy của đầu dò, do sự biến đổi của môi trường, do chính bản thân chất liệu sinh ra, do sai số lượng tử hóa hay sai số truyền v.v. Dựa trên tất cả các nguyên nhân gây ra nhiễu ở trên thì nhiễu sinh ra được phân thành các loại chính như sau:

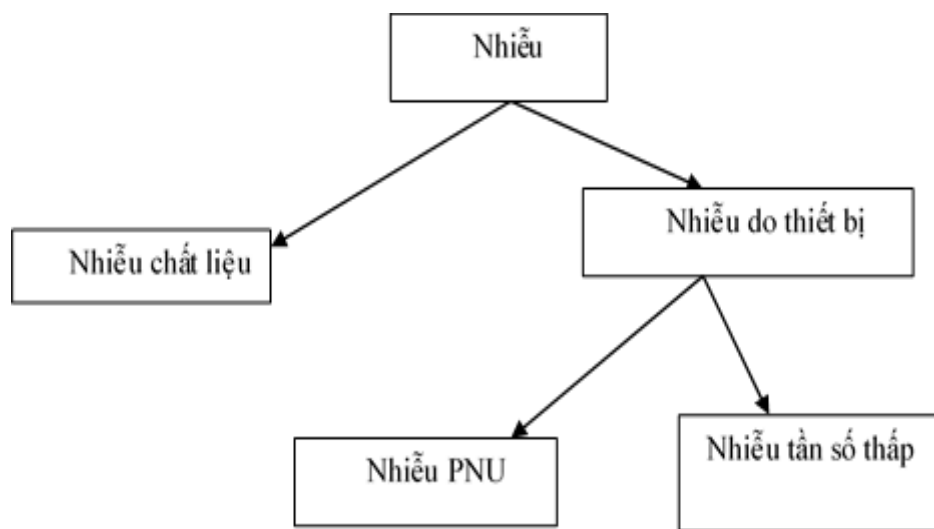
- Nhiễu độc lập với dữ liệu ảnh (Independent Noise): Là một loại nhiễu cộng (additive noise): Ảnh thu được  $f(i,j)$  là tổng của ảnh đúng (true image) và nhiễu  $n(i,j)$ :  $f(i,j) = s(i,j) + n(i,j)$ .
- Nhiễu phụ thuộc vào dữ liệu (Data dependent noise): Nhiễu xuất hiện khi có sự bức xạ đơn sắc nằm rải rác trên bề mặt ảnh, độ lớn chớm trên bề mặt tùy thuộc vào bước sóng của điểm ảnh. Do có sự giao thoa giữa các sóng ảnh nên làm xuất hiện những vết lốm đốm trên ảnh.
- 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) thuộc loại nhiễu cộng và độc lập (independent, additive noise)) – nhiễu  $n(i,j)$  có phân bố Gauss

(trung bình = 0) được mô tả bởi độ lệch chuẩn (standard deviation) hay phương sai. Mỗi pixel trong ảnh nhiễu là tổng giá trị pixel đúng (true pixel) 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 thiết lập luân phiên mang giá trị 0 hay giá trị cực đại tạo ra hình muối tiêu trên ảnh.

Thông thường trong các bài toán xử lý với ảnh số, người ta thường phải khử nhiễu trước khi sử dụng bức ảnh trong những ứng dụng nào đó, đặc biệt là trong các ứng dụng chuyên dụng của các lĩnh vực phân tích và xử lý hình ảnh. Gần đây, đã có một số ứng dụng của việc phân tích lượng nhiễu dư trong các bức ảnh để giải quyết bài toán phát hiện ảnh giả mạo dạng cắt dán, như nhóm nghiên cứu của Fridrich [40]. Nhóm đã xây dựng được nhiễu bất biến cho camera, với mỗi camera sẽ có một lượng nhiễu đặc trưng cho nó. Với một bức ảnh giả mạo dạng cắt dán cho trước, nhóm nghiên cứu này đã tính được ảnh nhiễu trong bức ảnh đó. Ảnh không phải là giả mạo nếu bức ảnh nhiễu của nó không tương thích tại tất cả các vùng. Dựa trên cơ sở phân tích nhiễu và đưa ra một mô hình cho nhiễu cảm biến, luận án ứng dụng khá thành công cho đề tài phát hiện ảnh giả mạo của Viện Khoa học và Công nghệ Việt Nam năm 2009 [28]. Tiếp đó, chúng tôi đồng thời cũng chứng minh được bằng thực nghiệm sự ảnh hưởng của lượng nhiễu này phụ thuộc vào từng chất liệu, mỗi chất liệu khác nhau sẽ có một độ đo nhiễu nhất định đặc trưng cho chất liệu đó. Dựa trên tư tưởng đó, luận án đã nghiên cứu và đề xuất một mô hình cho nhiễu chất liệu, từ mô hình này sẽ xây dựng các đặc trưng nhiễu bất biến cho chất liệu. Chúng tôi cũng sử dụng các đánh giá thống kê để tính toán đặc trưng nhiễu bất biến chất liệu, các đánh giá thống kê là phù hợp nhất cho việc mô tả và phát hiện chất liệu, đặc biệt là chất liệu tự nhiên. Có một vài nguồn không hoàn thiện và nhiễu đã có tác động đến bức ảnh trong quá trình xử lý. Khi camera thu nhận một bức ảnh từ môi trường, kết quả của ảnh vẫn có sự thay đổi nhỏ trong cường độ của các điểm ảnh riêng. Một phần nguyên nhân do các thành phần nhiễu ngẫu nhiên gây ra, cũng như nhiễu điện tử hoặc

nhiều chấm và một phần bởi do chính nhiều chất liệu hay nhiều đối tượng, chúng là thành phần xác định tồn tại dưới dạng xấp xỉ được sinh ra trong quá trình thu nhận thông qua bộ cảm biến. Với nhiễu do chất liệu sinh ra thực chất bị ảnh hưởng bởi ánh sáng tác động vào nó, ánh sáng này tác động lên mỗi chất liệu khác nhau sẽ sinh ra một lượng nhiễu khác nhau. Mặt khác, nhiễu do thiết bị bao gồm: nhiễu cố định (mẫu nhiễu có cường độ biến đổi trong miền biên độ ảnh) và nhiễu hỗn tạp không đều, điểm ảnh bất định (PNU: *pixel non-uniformity*) dựa trên sự khác nhau giữa pixel-pixel khi màn cảm biến không đặt vào nguồn sáng. Với nhiễu do thiết bị chúng ta có thể xấp xỉ nó thành một đặc trưng bất định. Hình 1.8 là một cách phân loại các loại nhiễu chất liệu. Do đó, sự khác nhau giữa chất liệu này với chất liệu khác có thể được xem xét thông qua nhiễu chất liệu. Xét tín hiệu thô  $x=(x_{i,j})$  với  $i=1,\dots,m, j=1,\dots,n$  với  $n, m$  là các kích thước của ảnh.



Hình 1.8. Sơ đồ phân loại nhiễu

Xét một tín hiệu nhiễu ngẫu nhiên:  $\eta = (\eta_{ij})$ , nhiễu được thêm vào do tác động bên ngoài là  $\varepsilon = (\varepsilon_{ij})$ , đốm đen hiện thời là  $c=(c_{ij})$ . Tín hiệu đầu ra được tính bởi  $y=(y_{ij})$  theo công thức sau:

$$y_{ij} = f_{ij}(x_{ij} + n_{ij} + \eta_{ij}) + c_{ij} + \varepsilon_{ij} \quad (1.3)$$



Đối với chất liệu, nhiều chất liệu  $n_{ij}$  là thành phần chủ yếu có trong nhiễu và nó thể hiện được sự phân biệt giữa các chất liệu khác nhau. Các thành phần còn lại là không đáng kể và có thể xem như một thành phần bất định.

Tín hiệu  $y$  thu được phải qua một chuỗi các xử lý phức tạp trước khi file ảnh cuối cùng được lưu trữ. Quá trình xử lý bao gồm các thao tác trên một lát giềng địa phương của các điểm ảnh, cũng như tái tạo lại màu, hiệu chỉnh màu, hoặc dùng các bộ lọc. Một vài toán tử không tuyến tính, cũng như hiệu chỉnh Gamma, tính toán số dư trắng, hoặc áp dụng nội suy màu. Điểm ảnh cuối cùng có giá trị là  $P_{ij}$ , với giả sử rằng  $0 \leq P_{ij} \leq 255$  cho mỗi kênh màu là :

$$P_{ij} = T(y_{ij}, N(y_{ij}), i, j) \quad (1.4)$$

ở đây  $T$  là một hàm không tuyến tính của  $y_{ij}$ , vị trí điểm ảnh  $(i, j)$ , và giá trị  $y$  từ một lát giềng địa phương  $N(y_{ij})$ .

### 1.2.3. Tiếp cận dựa vào mô hình

Tiếp cận này nhằm tìm ra các phương trình mô tả cho mẫu chất liệu. Tham số trong mô hình có thể mô tả các đặc tính bản chất của mẫu chất liệu. Các mô hình thường dùng để biểu diễn mẫu chất liệu như trường Markov ngẫu nhiên [22], Trường Gibb [47], mô hình Wold [21] đã được quan tâm nghiên cứu. Theo tiếp cận này, chất liệu được mô hình hóa bằng mô hình xác suất hoặc tổ hợp tuyến tính của một số hàm cơ sở. Các hệ số của mô hình chính là các biểu diễn cho mẫu chất liệu. Nhưng vấn đề quan trọng trong các phương pháp này là làm thế nào để ước lượng các tham số từ các mô hình đó và làm thế nào để chọn ra các mô hình chính xác phù hợp với từng mẫu chất liệu khác nhau. Các hệ số của các mô hình này thường được biến đổi thành các dạng khác nhau bất biến với phép dịch chuyển, quay, tỉ lệ hay ánh sáng v.v

#### 1.2.3.1. Mô hình SAR và RISAR

Mô hình SAR (Simultaneous Autoregressive Model) [117] đã được nhiều tác giả dùng trong phân lớp kết cấu bề mặt ảnh (texture).

Giả sử  $f(s)$  là giá trị mức xám của pixel  $s$  và mẫu chất liệu kích thước  $k \times l$ . Mô hình SAR được xác định như sau:

$$f(s) = u + \sum_{r \in \omega} \theta(r) f(s+r) + \varepsilon(s) \quad (1.5)$$

trong đó  $\omega$  là tập các lân cận của pixel  $s$ ,  $\varepsilon(s)$  là biến ngẫu nhiên Gauss độc lập,  $u$  là trung bình mức xám,  $\theta(r)$  là các tham số mô hình và chính là đặc trưng biểu diễn cho chất liệu. Mô hình SAR phụ thuộc phép quay nghĩa là  $\theta(r)$  thay đổi khi chất liệu bị quay. Để khắc phục nhược điểm phụ thuộc phép quay của mô hình SAR, Kasshyap Khotanzad đã phát triển SAR thành CSAR (Circular Simultaneous Autoregressive Model) [119]. Mô hình CSAR là phiên bản bất biến phép quay được xây dựng dựa trên lân cận hình tròn. Điều này có nghĩa chỉ những điểm xung quanh hình tròn của điểm trung tâm mới được sử dụng để miêu tả CSAR, do đó nó không miêu tả được chính xác mối quan hệ giữa một pixel và lân cận của nó. Mao và Jain [67] đã phát triển mô hình CSAR thành SAR bất biến đa biến với tên gọi là RISAR (Rotation Invariant SAR), phương pháp này xác định lân cận của một điểm bằng một số hình tròn xung quanh nó. Vì vậy khi ảnh bị quay xung quanh điểm này giá trị mức xám tại mỗi hình tròn xấp xỉ bằng nhau, và do đó sẽ bất biến với phép biến đổi quay. Các tham số mô hình xác định được chính là các đặc trưng bất biến quay.

### 1.2.3.2. Mô hình Markov

Cohen và các cộng sự [22] đã đề xuất trường Markov ngẫu nhiên Gauss và sử dụng phương pháp ước lượng hợp lý cực đại (Maximum Likelihood Estimate – MLE) để ước lượng các hệ số và các góc quay trong việc mô hình hóa kết cấu. Thực chất texture chính là một thuộc tính chất liệu, do đó nói khác hơn là có thể mô hình hóa mẫu chất liệu dựa vào trường Markov ngẫu nhiên. Khó khăn chính của phương pháp này là tính toán hàm hợp lý phi tuyến và cực đại địa phương, do đó cần phải dùng phương pháp lặp để tính toán sự nhạy cảm này. Chen và Kundu [21] đề xuất các mô hình cải tiến bất biến với phép quay dựa trên mô hình Markov ẩn (Hidden Markov Model – HMM) và phương pháp HMM đã được đánh giá hiệu quả trong nhiều nghiên cứu về phân lớp các kết cấu bề mặt.

### 1.2.3.3. Mô hình hình học Fractal

Fractal được biết và nghiên cứu bởi một số nhà toán học như Cantor, Poincare và Hilbert [13] từ cuối thế kỷ 19 đến đầu thế kỷ 20. Nhưng Mandelbrot [66] mới chính là người đặt nền móng toán học cho Fractal. Sau đó John Hutchinson [46] đã đề xuất lý thuyết hàm lặp ứng dụng cho nén ảnh Fractal, đây cũng là bước đột phá thứ hai của lý thuyết Fractal. Michael Barnsley [5] đã sử dụng lý thuyết hàm lặp để xây dựng định lý Collage mô tả cách thức mà một hệ thống hàm lặp sinh ra ảnh Fractal. Trong [47] Arnaud Jacquin là học trò của Barnsley đã cài đặt thuật toán chuyển đổi một bức ảnh thành hệ hàm lặp phân hoạch (Partitioned Iterated Function System-PIFS). Thuật toán này là cơ sở cho hầu hết các thuật toán mã hóa Fractal ngày nay. Mục tiêu của các thuật toán này thiết lập khả năng tạo ra một chuỗi các quá trình toán học cho phép tái tạo lại một bức ảnh chính xác. Nhiều thuật toán [72,111] cũng đề xuất sử dụng các mã này trong nén ảnh.

**Đặc điểm của Fractal:** Các hình dạng Fractal được đặc tả bởi sự tự tương tự về mặt thống kê, các quá trình đều đặn xuất hiện trên một dãy các tỉ lệ nhất định và chiều Fractal là không nguyên. Dù chúng ta thấy được các tính chất trực quan của nó trong rất nhiều ứng dụng nhưng sự hình tượng chúng rất khó và phức tạp được che giấu đằng sau các nghiên cứu Fractal cho đến khi có sự tiến bộ của khoa học máy tính gần đây, các ứng dụng Fractal mới dần được sáng tỏ hơn. Có thể tính toán chiều Fractal bằng nhiều phương pháp như đếm hộp (box-counting) đã ước lượng độ phức tạp của Fractal theo số các hộp sử dụng để xấp xỉ ảnh tại các mức tỉ lệ khác nhau [91]. Các biểu diễn ảnh Fractal thường có một số tính chất sau:

- **Không có chiều dài đặc tả:** Hình dạng nào cũng có một tỉ lệ xác định để đặc tả nó. Ví dụ hình dáng hình học có chiều dài đặc trưng riêng như bán kính hoặc chu vi của đường tròn, cạnh và đường chéo của hình vuông. Nhưng các hình Fractal không có các đơn vị đo như vậy. Chiều dài, kích thước hay chu vi của chúng không thể được xác định bằng các đơn vị riêng vì bề mặt của nó là không trơn làm cho người nhìn thấy gần hơn, các lỗ phức tạp trên bề

mặt cũng xuất hiện vì vậy chúng ta không thể vẽ được đường tiếp tuyến của các hình Fractal tức là không thể lấy đạo hàm tại điểm đó.

- **Tính tự tương tự:** Các hình Fractal là hình đơn vì không thể xác định bằng chiều dài đặc tả riêng do chúng ta có thể tìm ra các thuộc tính lặp lại tại các tỉ lệ khác nhau. Mặt khác, do hình Fractal thỏa mãn tính tự tương tự, nên hình dạng này không thay đổi thậm chí khi quan sát dưới các tỉ lệ khác nhau. Một ví dụ làm rõ điều này nhất đó là hình của lá xà cừ.
- **Chiều Fractal không nguyên và lớn hơn chiều tô pô:** Giả sử ta xem 1 điểm có chiều topo là 0, đường có chiều tô pô là 1, bề mặt có chiều tô pô là 2 và hình trụ có chiều tô pô là 3. Tuy nhiên một đường cong phức trên bề mặt có chiều Fractal là một số thực giữa 1 và 2, đường cong tiếp giáp bề mặt có chiều xấp xỉ là 2. Giá trị thật sự của chiều Fractal cũng hơi khác phụ thuộc vào phương pháp xác định, gần đây cũng có một số phương pháp khả thi về mặt vật lý. Chúng ta có thể đánh giá chiều Fractal bằng sự thay đổi mức thô-mịn (phương pháp đếm hộp), dựa trên quan hệ độ đo Fractal, sử dụng hàm tương quan, sử dụng hàm phân phối hoặc dựa trên phổ năng lượng.
- **Không gian metric:** Một không gian  $\mathcal{M}$  (có thể là không gian con compact của  $\mathbb{R}^3$ ) là một không gian metric nếu với bất kỳ 2 phần tử  $x, y$  thì tồn tại số thực  $d(x,y)$  gọi là khoảng cách từ  $x$  đến  $y$  thỏa mãn các tính chất sau:

$$(1) d(x, y) \geq 0$$

$$(2) d(x, y) = 0 \Leftrightarrow x = y$$

$$(3) d(x, y) = d(y, x)$$

$$(4) d(x, z) \leq d(x, y) + d(y, z)$$

- ✓ **Chuỗi Cauchy:** Một chuỗi  $\{x_n\}_{n=0}^{\infty} = \{x_n \in \mathcal{M}, n \in \mathbb{N}\}$  được gọi là chuỗi Cauchy nếu  $\forall \varepsilon > 0, \exists K \in \mathbb{N}$  sao cho  $d(x_n, x_m) \leq \varepsilon, \forall n, m > K$

- ✓ Không gian metric đầy đủ: Một không gian metric  $(\mathcal{M}, d)$  là đầy đủ nếu với mọi chuỗi Cauchy của các điểm  $\{x_n\}_{n=0}^{\infty}$  trong  $\mathcal{M}$  có giới hạn  $x_n \in \mathcal{M}$ .
- ✓ **Ánh xạ co**: Một phép biến đổi  $w: \mathcal{M} \rightarrow \mathcal{M}$  được gọi là ánh xạ co với hệ số co  $s \in [0,1)$  nếu với bất kỳ hai điểm  $x, y \in \mathcal{M}$  thì khoảng cách:

$$d(w(x), w(y)) < s \cdot d(x, y)$$

Điều này muốn nói lên rằng ánh xạ co luôn luôn mang các điểm lại gần nhau hơn (vì hệ số co nhỏ hơn 1).

Ánh xạ co có tính chất đó là khi áp dụng nhiều lần liên tiếp, thì chúng sẽ hội tụ đến một điểm bất động duy nhất.

**Định lý điểm bất động [5]**: Nếu  $(\mathcal{M}, d)$  là không gian metric đầy đủ và  $w: \mathcal{M} \rightarrow \mathcal{M}$  là ánh xạ co với hệ số co  $s$ , thì:

- ✓ Tồn tại một điểm bất động duy nhất  $x_f \in \mathcal{M}$  là bất biến đối với  $w$ :

$$w(x_f) = x_f$$

- ✓ Với  $x \in \mathcal{M}$  bất kỳ, ta luôn có:

$$\lim_{n \rightarrow \infty} w^n(x) = \lim_{n \rightarrow \infty} w(w(\dots(x))) = x_f$$

- ✓ Định lý Collage: Với bất kỳ  $x \in \mathcal{M}$  ta có:

$$d(x, x_f) \leq \frac{1}{1-s} d(x, w(x))$$

Định lý điểm bất động chứng minh được mã hóa Fractal của một bức ảnh bất kỳ là khả thi. Xét các ảnh như là các điểm trong không gian metric và tìm ánh xạ co trên không gian này mà điểm bất động của nó chính là ảnh cần mã hóa (trong thực tế đó chính là ảnh gần với ảnh cần mã hóa). Định lý điểm bất động cũng đảm bảo khoảng cách giữa điểm được biến đổi (bằng ánh xạ co đó) và điểm bất động nhỏ hơn khoảng cách giữa điểm ban đầu và điểm bất động. Nếu chúng ta áp dụng nhiều lần liên tiếp

ánh xạ co lên điểm ban đầu thì chúng ta sẽ nhận được các điểm càng gần với điểm bất động.

**Phép biến đổi afin** : Với ảnh  $J$  cho trước, nếu ta ký hiệu  $z$  là cường độ pixel tại vị trí  $(x,y)$  thì phép biến đổi afin  $W$  có thể được biến đổi dưới dạng ma trận như sau:

$$W \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & s \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} e \\ f \\ o \end{bmatrix}$$

Trong đó  $a, b, c, d, e, f$  là các tham số hình học và  $s$  là độ tương phản,  $o$  là độ sáng, phép biến đổi này có thể viết lại dưới dạng tuyến tính  $W(X) = AX + B$ ,  $A$  là ma trận  $n \times n$  (trong trường hợp này  $n = 3$ ),  $B$  là vectơ kích thước  $n \times 1$ . Phép biến đổi afin này ta có thể tính toán ước lượng được các tham số của phép tỉ lệ, phép quay ảnh hoặc dịch chuyển trên cường độ điểm ảnh.

### **Hệ hàm lặp (Iterated Function Systems - IFS):**

Một hệ hàm lặp  $\{W: w_i, i = 1, 2, \dots, N\}$  là một tập các phép biến đổi afin co  $w_i: \mathcal{M} \rightarrow \mathcal{M}$  với các hệ số co  $s_i$  trên không gian metric đầy đủ  $(\mathcal{M}, d)$ , tập các ánh xạ này xác định một phép biến đổi co  $W$  với hệ số co  $s = \max\{s_i, i = 1, 2, \dots, N\}$ . Phép biến đổi co  $W$  trên không gian metric đầy đủ  $(\mathcal{M}, d)$  sẽ có một điểm cố định  $X_f$  cũng được gọi là nhân của IFS:

$$W(X) = \bigcup_{i=1}^N w_i(X)$$

$$W(X_f) = \bigcup_{i=1}^N w_i(X_f) = X_f$$

**Nguyên tắc mã hóa Fractal**: Có nhiều phương pháp mã hóa Fractal đề xuất, chúng chỉ khác phương pháp phân hoạch, lớp phép biến đổi, kiểu tìm kiếm để xác định vị trí của khối miền phù hợp. Thuật toán mã hóa ảnh tự động đầy đủ đầu tiên được đưa ra bởi Jacquin [47] năm 1989 cho đến khi bộ mã hóa của Jacquin trở nên

phổ biến. Các nhà nghiên cứu đã cố gắng thiết kế các mã hóa Fractal hỗ trợ tạo ra các phép biến đổi bằng cấu trúc hàm lặp. Phương pháp của Jacquin dựa trên hệ hàm lặp *PIFS* là một cải tiến mới của IFS, ở đây các phép biến đổi tổng quát hơn khảo sát thực tế rằng một phần của ảnh có thể được xấp xỉ bởi sự biến đổi và lấy mẫu lại của phần khác trong cùng ảnh, tính chất này được gọi là tự tương tự từng phần. Một *PIFS* gồm một không gian metric đầy đủ  $\mathcal{M}$ , tập các miền  $D_i \in \mathcal{M}, i = 1, 2, \dots, n$  và tập các ánh xạ co  $w_i: D_i \rightarrow \mathcal{M}, i = 1, \dots, n$ . Phương pháp này mô tả như sau:

- ✓ Xây dựng khối dĩa: Ảnh cần mã hóa được phân hoạch thành các khối dĩa không chồng lên nhau  $R_i$
- ✓ Xây dựng khối miền: Phân hoạch ảnh thành các khối miền  $D_j$  lớn hơn nên có thể chồng nhau.
- ✓ Biến đổi: Nhiệm vụ của biểu diễn Fractal cho mẫu chất liệu là phải tìm ra khối miền  $D_{R_i}$  trong cùng ảnh cho mỗi  $R_i$  sao cho mỗi phiên bản biến đổi của  $w(D_{R_i})$  là một xấp xỉ tốt của  $R_i$ , phép biến đổi co  $w$  là tổ hợp của phép biến đổi hình học và phép biến đổi ánh sáng. Phiên bản được biến đổi của khối miền có thể được quay, thu nhỏ, lấy tỉ lệ tương phản và dịch chuyển. Vì vậy phép biến đổi này phải là phép biến đổi afin.

Phân hoạch: Bước quyết định đầu tiên trong các phương pháp mã hóa ảnh đó là việc lựa chọn kiểu phân hoạch ảnh thành khối miền và khối dĩa. Phân hoạch khối dĩa đơn giản nhất là chia ảnh thành các khối hình vuông kích thước cố định. Phân hoạch cây tứ phân là một kỹ thuật trong xử lý ảnh dựa trên việc tách đệ quy các nhánh được lựa chọn, vì vậy kết quả phân hoạch cho ta một cấu trúc cây, trong đó mỗi nút không kết thúc có bốn nhánh con.

Phân hoạch ngang-dọc cũng giống phân hoạch cây tứ phân, tạo ra phân hoạch cấu trúc cây của ảnh. Tuy nhiên thay vì tách đệ quy các nhánh thì mỗi khối được tách thành hai theo đường ngang, đường dọc và cuối cùng một số các kiểu phân hoạch như phân hoạch tam giác trong đó một ảnh hình chữ nhật được chia thành hai tam

giác theo đường chéo. Mỗi trong số đó lại được chia nhỏ thành bốn tam giác bằng tách tam giác này theo đường nối ba điểm phân hoạch dọc theo cạnh của tam giác.

Phép biến đổi: Một bước quan trọng trong việc biểu diễn Fractal cho mẫu chất liệu đó là lựa chọn loại phép biến đổi vì nó quyết định thuộc tính hội tụ của giải mã và các tham số định tính về thông tin nội dung ảnh. Định lý điểm bất động yêu cầu phép biến đổi phải co, thông qua điểm cố định của chúng để biểu diễn các điểm trong không gian. Tuy nhiên định lý này không cho ta cách tìm các phép biến đổi như vậy.

Nếu chúng ta tìm được một ánh xạ co  $W$  hợp lý cho ảnh  $X_f$ . Chúng ta biết điểm bất động của  $W$  là  $X_f$ , nên:

$$d(X_f, W(X_f)) = d(X_f, X_f) = 0$$

Thật sự là rất khó tìm ra phép biến đổi  $W$  một cách chính xác cho ảnh  $X$  bất kỳ nào. Thay vì vậy, nhiều phương pháp mã hóa ảnh chỉ tìm một phép biến đổi  $W^*$  với nhân  $X_f^*$  với  $d(X, X_f^*)$  nhỏ nhất có thể. Nếu khoảng cách  $d(X, W(X)) \leq \delta$  thì khoảng cách từ  $X$  đến xấp xỉ  $X_f^*$  có thể được lấy cận dưới như sau:

$$d(X, X_f^*) \leq \frac{\delta}{1 - s}$$

Do đó  $s$  và  $\delta$  phải nhỏ nhất có thể. Phép biến đổi affine là tốt trong trường hợp như vậy. Mỗi phép biến đổi có hai phần: hình học và ánh sáng. Phần hình học các phép biến đổi gồm: phép lấy tỉ lệ, quay, dịch chuyển một block miền để làm vừa với block dãy. Để thỏa mãn phép biến đổi là co, kích thước của block miền phải luôn luôn lớn hơn block dãy vì vậy hệ số tỉ lệ  $s$  luôn nhỏ hơn 1. Phần ánh sáng bao gồm một số các phép đơn giản dịch chuyển, tỉ lệ độ tương phản.

Phát hiện mẫu chất liệu dựa vào hình học Fractal là một vấn đề mới và khó nhưng rất hiệu quả do bản chất tự lặp lại các chi tiết ở những tỉ lệ khác nhau, do đó chúng giải quyết được vấn đề thay đổi tỉ lệ toàn cục mà hiện nay các nhà nghiên cứu hết sức quan tâm. Có một vài nghiên cứu gần đây về xác thực kí tự trong an ninh cũng dựa trên tiếp cận IFS của hình học Fractal và được đánh giá cao trong khoa học thị



giác máy. Ý tưởng Fractal cũng có một số nghiên cứu trên chất liệu nhưng chỉ dừng lại ở mức ước lượng chiều Fractal và chỉ phù hợp cho ảnh mức xám.

### **1.3. Kết luận và vấn đề nghiên cứu**

Phát hiện chất liệu hay phát hiện mẫu chất liệu trong ảnh là một vấn đề hết sức quan trọng trong thị giác máy. Trong chương này, luận án đã trình bày các khái niệm về chất liệu, mẫu chất liệu và bài toán phát hiện mẫu chất liệu trong ảnh cũng như những thách thức đặt ra trong thực tế của bài toán. Trên cơ sở đó, hệ thống hóa các nghiên cứu liên quan theo các cách tiếp cận biểu diễn và phát hiện chất liệu nhằm mục đích xây dựng một số thuật toán phát hiện mẫu chất liệu trong ảnh để khắc phục những trường hợp có sự biến đổi hình học và quang học trong quá trình thu nhận ảnh như phép thay đổi tỉ lệ, thay đổi quay, dịch chuyển và thay đổi ánh sáng. Đây cũng chính là mục tiêu nghiên cứu trong các phần tiếp theo của luận án. Đồng thời với đó là nghiên cứu ứng dụng của kỹ thuật đề xuất vào một số bài toán có tính thời sự như bài toán phát hiện ảnh số giả mạo, phát hiện vào ra trong giám sát tự động v.v.

## **Chương 2. PHÁT HIỆN MẪU CHẤT LIỆU DỰA VÀO ĐẶC TRƯNG BẤT BIẾN ĐỊA PHƯƠNG**

### **2.1. Đặt vấn đề**

Đặc trưng bất biến địa phương đã được nghiên cứu nhiều cả về lý thuyết lẫn ứng dụng trong thời gian qua, nhằm mục đích đối sánh các biến thể ảnh do điều kiện thu nhận khác nhau của cùng một đối tượng hoặc một mẫu chất liệu. Một trong những đặc trưng bất biến địa phương được đánh giá cao đó là đặc trưng bất biến tỉ lệ hay còn gọi là SIFT (Scale Invariant Feature Transform) [60]. Đặc trưng này có tính phân biệt cao, bất biến đối với các phép biến đổi như: tỉ lệ (scale), phép quay (rotation) và ít ảnh hưởng do các phép biến đổi afin, thay đổi kiểu chụp 3D, nhiễu và cả sự thay đổi ánh sáng nhẹ. Thực chất SIFT chính là phép biến đổi dữ liệu ảnh vào hệ tọa độ bất biến tỉ lệ. Năm 2006 C. Schmid và các cộng sự [55,70] đã có nghiên cứu đánh giá các đặc trưng địa phương khác nhau như Hessian & Harris [44], Hessian/Harris – Laplace/ Afin, SIFT, LBP, CS-LBP trong nhận dạng, tra cứu ảnh và cho thấy rằng SIFT là đặc trưng tiềm năng nhất trong trường hợp nhận dạng cảnh hay đối tượng và có ưu điểm tính toán nhanh, hiệu suất thời gian thực. Gần đây, cũng có nhiều nghiên cứu nhận dạng chất liệu, nhận dạng kết cấu dựa vào đặc trưng bất biến địa phương SIFT [53,55,59] nhưng tất cả đều sử dụng đối sánh trực tiếp các đặc trưng SIFT, cho đến nay đặc trưng bất biến địa phương vẫn được sử dụng trong nhiều nghiên cứu gần đây [90,116,118]. Trên cơ sở đó, luận án đề xuất kỹ thuật sử dụng đặc trưng bất biến địa phương cho phát hiện mẫu chất liệu dựa trên việc tìm kiếm mẫu chất liệu trong ảnh trên cơ sở tìm kiếm cấu trúc tương quan hình học các đặc trưng của mẫu chất liệu với các đặc trưng trong ảnh cần phát hiện. Kỹ thuật đề xuất được cài đặt và đánh giá trên tập cơ sở dữ liệu ảnh trích ra từ tập dữ liệu video phát hiện và phân tích hoạt động của đám đông trên đường phố MIT - CSAIL cùng một số ảnh thu nhận được với các mẫu chất liệu trong tự nhiên. Kết quả cho thấy kỹ thuật đề xuất phát hiện khá chính xác các vùng chứa mẫu chất liệu nhờ sử dụng cấu trúc tương quan hình học, thậm chí khi mẫu chất liệu bị thay đổi ánh sáng nhẹ, thay đổi tỉ lệ và bị quay.

Mặt khác, phát hiện ảnh số giả mạo hiện đang là một vấn đề hết sức cấp bách trong tình hình kinh tế xã hội đang sôi động của mỗi quốc gia, đặc biệt là Việt Nam. Nhiều quyết định phải dựa vào tính thật của ảnh số như trong an ninh, y tế v.v. Đã có nhiều nghiên cứu về bài toán phát hiện ảnh giả mạo trong những năm gần đây, nhưng vẫn tồn tại nhiều khó khăn như việc phát hiện vùng nghi ngờ giả mạo khi có sự thay đổi về phép quay, phép thay đổi tỉ lệ vẫn chưa thực hiện được [40,41]. Ở chương này, luận án cũng đề xuất phương pháp giải quyết bài toán phát hiện ảnh số giả mạo dạng cắt dán trong trường hợp vùng nghi ngờ giả mạo bị thay đổi quay và thay đổi tỉ lệ dựa trên tiếp cận phát hiện mẫu chất liệu sử dụng đặc trưng bất biến địa phương trên cơ sở cải tiến thuật toán Exact Match\* [28]. Đề xuất này được cài đặt và kết quả cho thấy đã khắc phục được khó khăn của các thuật toán phát hiện ảnh số giả mạo hiện nay.

## **2.2. Trích chọn đặc trưng bất biến địa phương cho mẫu chất liệu**

Như phân tích trong chương 1, việc phát hiện mẫu chất liệu dưới những điều kiện thu nhận ảnh không biết trước thì sự thay đổi tỉ lệ làm ảnh hưởng lớn đến quá trình phát hiện [19]. Do đó, áp dụng đặc trưng bất biến địa phương trong mô tả mẫu chất liệu, luận án sử dụng đặc trưng bất biến tỉ lệ. Phát hiện các đặc trưng bất biến tỉ lệ được đề xuất gồm 2 pha chính:

- Tìm các điểm bất biến tỉ lệ bằng cách xác định các điểm cực trị địa phương trên không gian tỉ lệ.
- Xây dựng mô tả cho các điểm bất biến tỉ lệ đã chọn để thu được đặc trưng biểu diễn mẫu chất liệu.

### **2.2.1. Tìm các điểm bất biến địa phương trên không gian tỉ lệ**

David Lowe [60] sử dụng cách tiếp cận lọc theo tầng để xác định các điểm bất biến tỉ lệ trên từng tầng, mỗi tầng chính là một ảnh với một tỉ lệ cụ thể và chỉ các điểm này mới được khảo sát tiếp theo. Bước đầu tiên để tìm điểm bất biến tỉ lệ là phải xác định vị trí và tỉ lệ tại các hướng chụp khác nhau của cùng một đối tượng. Phát

hiện các điểm bất biến tỉ lệ này bằng cách tìm các điểm ổn định trên tất cả các tỉ lệ dựa trên hàm tỉ lệ liên tục giống như một không gian tỉ lệ có dạng  $f = kernel * I(x, y)$  (Witkin, 1983). Koenderink (1984) và Lindeberg (1994) đã chứng minh trong nhiều trường hợp giả thuyết chỉ có kernel Gauss là phù hợp nhất. Vì vậy, có thể biểu diễn ảnh trên không gian tỉ lệ bằng hàm  $L(x, y, \sigma)$ , trong đó:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$

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

Để phát hiện được vị trí của các điểm bất biến tỉ lệ ổn định trong không gian tỉ lệ, David Lowe trong [60] đã đề xuất tìm cực trị trên không gian tỉ lệ của hàm  $D(x, y, \sigma)$  (Difference – of – Gauss) chính là hiệu của hai ảnh được làm trơn Gauss có tỉ lệ gần nhau phân biệt bằng hệ số nhân  $k$ :

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned}$$

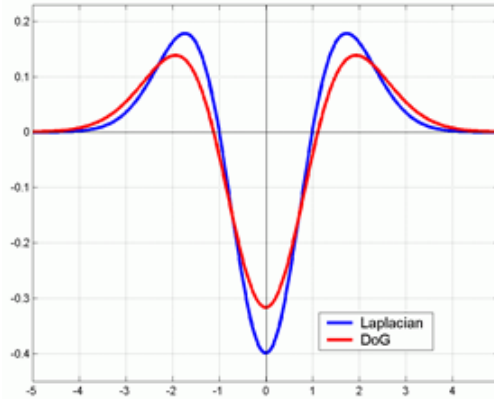
Sử dụng hàm  $D(x, y, \sigma)$  là một lựa chọn tốt vì nó có ưu điểm tính toán đơn giản bằng cách chỉ thực hiện phép toán trừ các ảnh được cuộn với hàm Gauss tại các tỉ lệ liên tiếp. Hơn nữa, hàm  $D(x, y, \sigma)$  là một xấp xỉ của hàm Laplace Gauss chuẩn hóa tỉ lệ  $\sigma^2 \nabla^2 G$  được Lindeberg nghiên cứu năm 1994. Ông đã chứng minh hàm Laplace- Gauss chuẩn hóa ở trên với hệ số  $\sigma^2$  là bất biến tỉ lệ đúng. Trong đánh giá thực nghiệm Mikolajczyk năm 2005 [70] nhận xét rằng cực trị của  $\sigma^2 \nabla^2 G$  cho ta các đặc trưng ổn định nhất so với sử dụng các hàm có thể khác. Mối quan hệ giữa  $D(x, y, \sigma)$  và  $\sigma^2 \nabla^2 G$  thể hiện thông qua phương trình nhiệt như sau:

$$\frac{\partial G}{\partial \sigma} = \sigma \nabla^2 G$$

Suy ra: 
$$\sigma \nabla^2 G = \frac{\partial G}{\partial \sigma} \approx \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma - \sigma}$$

Nghĩa là: 
$$G(x, y, k\sigma) - G(x, y, \sigma) = (k - 1)\sigma^2 \nabla^2 G$$

Từ đó có thể thấy rằng cực trị trên không gian tỉ lệ  $L(x, y, \sigma)$  chính là cực trị trên hàm  $D(x, y, \sigma)$ . Hình 2.1 minh họa  $D(x, y, \sigma)$  là xấp xỉ của  $L(x, y, \sigma)$ .



Hình 2.1.  $D(x,y,\sigma)$  xấp xỉ với  $L(x,y,\sigma)$

Để xác định được các điểm bất biến tỉ lệ, chỉ cần tìm các cực trị trên các ảnh  $D(x, y, \sigma)$ . Vấn đề thứ nhất ta phải chọn tần số lấy tỉ lệ (tần số trên không gian tỉ lệ) là bao nhiêu, tức là tìm cực trị trên bao nhiêu ảnh  $D$  liên tiếp với hệ số  $k$  như thế nào. Vấn đề thứ hai là phải chọn tần số lấy mẫu trên miền không gian ảnh, nghĩa là cần xác định  $\sigma$  bé nhất trong hàm Gauss là bao nhiêu. Qua đánh giá thực nghiệm trên một tập các ảnh của các quang cảnh là ổn định nhất so với các nghiên cứu khác và xét trên tất cả các phép biến đổi ảnh của David Lowe năm 2004, cho thấy tần số tỉ lệ được chọn là 3, nghĩa là ta chỉ xét trên 3 ảnh  $D$  liên tiếp với hệ số  $k = \sqrt{2}$  và  $\sigma$  bé nhất được khởi tạo trong hàm Gauss là 1.6.

Như vậy, việc tìm các điểm bất biến tỉ lệ được thực hiện bằng cách so sánh mỗi điểm của ảnh  $D$  trên mỗi tỉ lệ với 8 điểm lân cận trên cùng cùng tỉ lệ đó và với 18 điểm lân cận trên 2 tỉ lệ trên và dưới của ảnh  $D$ . Kết thúc bước này ta được tập các điểm bất biến tỉ lệ ứng cử. Các điểm được xét chỉ là điểm lấy mẫu và thường thì người ta lấy điểm trung tâm làm điểm bất biến tỉ lệ ứng cử. Bây giờ chúng ta phải xác định chính xác lại vị trí và tỉ lệ, đồng thời loại bỏ các điểm bất biến tỉ lệ có độ tương phản thấp và các điểm bất biến tỉ lệ dọc biên.

- **Xác định chính xác vị trí các điểm bất biến tỉ lệ:** Năm 2002 Brown và Lowe [17] đã đề xuất kỹ thuật nội suy để xác định chính xác vị trí điểm bất biến tỉ lệ bằng cách sử dụng khai triển Taylor của hàm không gian tỉ lệ  $D(x, y, \sigma)$  với gốc tại vị trí điểm mẫu  $X = (x, y, \sigma)$  :

$$D(X) = D + \frac{\partial D^T}{\partial X} X + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X \quad (2.1)$$

$D$  là giá trị của hàm tại điểm  $(0,0,0)$  và các giá trị đạo hàm của nó được lấy tại điểm mẫu,  $X = (x, y, \sigma)$  là độ lệch so với điểm đó. Xác định cực trị  $\hat{X}$  của đạo hàm  $D(X)$  bằng cách rút đạo hàm bậc nhất từ phương trình (2.1) và cho nó bằng 0. Giải phương trình này ta được:

$$\hat{X} = -\frac{\partial^2 D^{-1}}{\partial X^2} \frac{\partial D}{\partial X} \quad (2.2)$$

Thực chất đạo hàm của  $D$  cũng được xác định bằng cách lấy hiệu các điểm mẫu lân cận  $3 \times 3$ , nên chi phí tính toán của giai đoạn này cũng nhỏ. Nếu  $\hat{X} > 0.5$  theo mỗi hướng, nghĩa là nó gần với các điểm mẫu khác đã được chọn, ta thực hiện nội suy để thay cho điểm đó.

- **Loại bỏ những điểm bất biến tỉ lệ có độ tương phản thấp :** Để loại bỏ các điểm có độ tương phản thấp ta phải kiểm chứng bằng giá trị của  $D(\hat{X})$  bằng cách thay (2.2) vào phương trình (2.1) ta được:

$$D(\hat{X}) = D + \frac{1}{2} \frac{\partial D^T}{\partial X} \hat{X} \quad (2.3)$$

Theo thực nghiệm, điểm mẫu bị bỏ khi  $|D(\hat{X})| \leq 0.03$

- **Loại bỏ các điểm bất biến tỉ lệ dọc theo biên:**

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$$

Sử dụng cách tiếp cận của Harris và Stephens [42]. Ta xét đường cong chính của ảnh bằng ma trận Hessian:

Các đạo hàm này được tính bằng cách lấy hiệu các điểm mẫu lân cận.

Sau đó tìm được các trị riêng của  $H$  là  $\alpha$ ,  $\beta$ .

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

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

$$\alpha = r\beta$$

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

Loại bỏ điểm bất biến tỉ lệ chính xác xác định được nếu thỏa mãn bất đẳng thức:

$$\frac{Tr(\mathbf{H})^2}{Det(\mathbf{H})} < \frac{(r+1)^2}{r} \quad (2.4)$$

với  $r$  chọn được theo thực nghiệm là  $r = 10$ .

Sau một quá trình nhiều bước chúng ta chọn ra được vị trí chính xác của các điểm bất biến tỉ lệ ổn định.

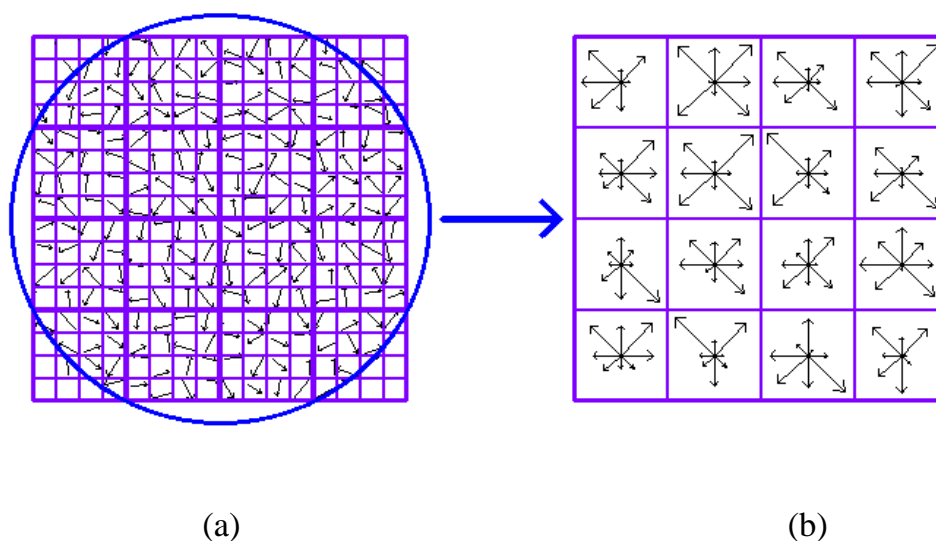
### 2.2.2. Xây dựng mô tả địa phương

Để các điểm bất biến tỉ lệ đã xác định bất biến được với phép quay và một phép biến đổi khác do ảnh hưởng của hướng thu nhận khác nhau. Trước hết, một cách hiệu quả nhất đó là sử dụng hướng và độ lớn gradient cho mỗi điểm mẫu. Mỗi điểm mẫu xác định bằng vị trí và tỉ lệ, chúng ta có thể gán hướng  $\theta(x, y)$  và độ lớn gradient  $m(x, y)$  như sau:

$$\theta(x, y) = \tan^{-1}((L(x, y + 1) - L(x, y - 1))/(L(x + 1, y) - L(x - 1, y))) \quad (2.5)$$

$$m(x, y) = \sqrt{(L(x + 1, y) - L(x - 1, y))^2 + (L(x, y + 1) - L(x, y - 1))^2} \quad (2.6)$$

Sau đó với mỗi điểm bất biến tỉ lệ đã được xác định trong phần 2.2.1, chúng ta xây dựng mô tả bằng cách tính toán hướng và độ lớn gradient của mỗi điểm mẫu xung quanh vị trí điểm bất biến tỉ lệ đó theo (2.5) và (2.6) bằng cách tính toán các lược đồ hướng cho mỗi vùng con  $4 \times 4$  lân cận theo 8 hướng. Mô tả đặc trưng được xây dựng bằng cách tùy chọn các lân cận điểm bất biến tỉ lệ theo các lược đồ hướng các vùng con. Thông thường, mô tả cho các điểm bất biến tỉ lệ bằng các lược đồ  $4 \times 4$  nghĩa là sử dụng lân cận  $16 \times 16$  điểm mẫu. Do vậy, đặc trưng thường dùng là một mô tả  $4 \times 4 \times 8 = 128$  chiều như hình 2.2. Sau bước này ta thu được một tập các vectơ mô tả cho chất liệu, bất biến với tỉ lệ, quay và sự thay đổi ánh sáng. Tiếp theo, luận án trình bày kỹ thuật phát hiện chất liệu dựa trên tập các mô tả đặc trưng bất biến đã tìm được này.



Hình 2.2. Xây dựng mô tả cho điểm bất biến tỉ lệ

(a) Ảnh gradient; (b) Mô tả cho điểm bất biến tỉ lệ

### 2.3. Phát hiện mẫu chất liệu dựa vào đặc trưng bất biến địa phương

Tìm mẫu chất liệu trên ảnh gồm 3 pha: Đối sánh để tìm ra các đặc trưng trên mẫu chất liệu tương ứng với đặc trưng trên ảnh. Sau bước này ta sẽ xác định được 2 tập đặc trưng tương ứng trên mẫu chất liệu và trên ảnh, nhưng sự tương ứng này



không có gì đảm bảo là chính xác. Do đó pha thứ 2, có thể kiểm chứng sự tương ứng này dựa trên phương pháp hình học bằng cách tìm điểm trung tâm ứng cử trên ảnh dựa vào điểm trung tâm trên mẫu chất liệu. Nếu mẫu chất liệu tồn tại trên ảnh thì các điểm trung tâm này phải tập trung tại cùng vị trí. Vì vậy pha thứ 3 sẽ phân cụm tập các điểm ứng cử trung tâm trên ảnh dựa trên ngưỡng khoảng cách và nếu số điểm trong mỗi cụm lớn hơn một ngưỡng cho trước thì kết luận mẫu chất liệu có tồn tại trên ảnh. Các pha được trình bày cụ thể như sau:

### 2.3.1. Đối sánh dựa vào phương pháp lân cận gần nhất

Trước hết, trích chọn các đặc trưng bất biến tỉ lệ cho ảnh vào  $\mathcal{J}$  và mẫu chất liệu  $\mathcal{M}$ . Mỗi đặc trưng được xác định bởi vị trí, tỉ lệ và giá trị gradient. Với mỗi đặc trưng của mẫu chất liệu  $\mathcal{M}$ , ta tìm đặc trưng tương ứng giống với nó nhất trên ảnh vào  $\mathcal{J}$  dựa trên khoảng cách Euclid, tức là giả sử gọi  $F_M^i$  là đặc trưng thứ  $i$  của mẫu chất liệu  $\mathcal{M}$ ,  $F_I^j$  là đặc trưng thứ  $j$  trên ảnh  $\mathcal{J}$ . Tìm đặc trưng  $F_I^{j'}$  trên  $\mathcal{J}$  thỏa mãn hàm khoảng cách Euclid đạt cực tiểu như sau:

$$j' = \arg \min_j \sqrt{\sum_{k=1}^{128} \left( (F_M^i)_k - (F_I^j)_k \right)^2} \quad (2.7)$$

Sau quá trình này ta có được tập các cặp đặc trưng của chất liệu  $\mathcal{M}$  tương ứng với tập các đặc trưng trên ảnh  $\mathcal{J}$  bằng phương pháp lân cận gần nhất. Bây giờ để xác định được chất liệu  $\mathcal{M}$  có trên ảnh vào  $\mathcal{J}$  hay không ta phải thực hiện quá trình kiểm chứng dựa vào cấu trúc hình học của các đặc trưng để xác định lại sự tương ứng giữa các cặp đặc trưng đó là chính xác.

### 2.3.2. Xác định sự tương ứng của mẫu chất liệu trong ảnh

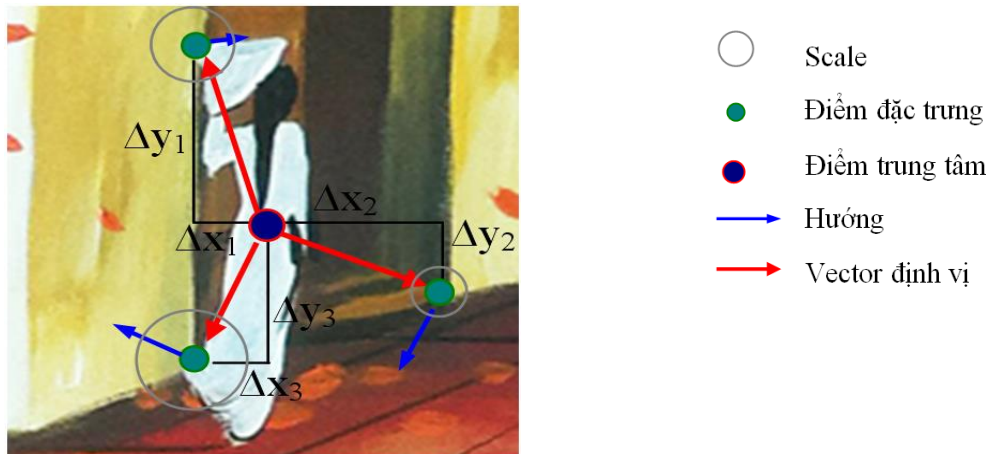
Chọn điểm ở giữa của mẫu chất liệu  $\mathcal{M}$  làm điểm trung tâm. Với mỗi đặc trưng của mẫu chất liệu  $\mathcal{M}$ , xác định được vectơ định vị  $\Delta = (\Delta_x, \Delta_y)$  như hình 2.3. Dựa vào vectơ định vị của mỗi đặc trưng này, tính toán điểm ứng cử trung tâm của mẫu chất liệu  $\mathcal{M}$  trên ảnh vào  $\mathcal{J}$  bằng các biểu thức sau:

$$X = x_I + \frac{\sigma_I}{\sigma_M} \times \sqrt{(\Delta_x^2 + \Delta_y^2)} \times \cos(\theta + \theta_M - \theta_I) \quad (2.8)$$

$$Y = y + \frac{\sigma_I}{\sigma_M} \times \sqrt{(\Delta_x^2 + \Delta_y^2)} \times \sin(\theta + \theta_M - \theta_I)$$

Trong đó  $\theta = \arctan(\frac{\Delta y}{\Delta x})$ ;  $\sigma_I, \sigma_M$ ;  $\theta_I, \theta_M$  là tỉ lệ và hướng của cặp đặc trưng tương ứng và  $x_I, y_I$  là tọa độ của điểm đặc trưng tương ứng trên ảnh vào  $\mathcal{J}$ .

Như vậy, ta tính được một tập các điểm ứng cử trung tâm trên ảnh vào  $\mathcal{J}$ . Khi mẫu chất liệu  $\mathcal{M}$  tồn tại trong ảnh  $\mathcal{J}$  thì các điểm ứng cử làm trung tâm này phải giống nhau (vị trí giống nhau). Nhưng thực chất các điểm ứng cử trung tâm tính được này không giống nhau hoàn toàn mà có thể là các vị trí lân cận gần nhau. Do đó ta cần phải phân cụm các điểm ứng cử trung tâm thành một số cụm để xác định mẫu chất liệu trên ảnh.



Hình 2.3. Cách xác định vec tơ định vị

### 2.3.3. Phân cụm các điểm ứng cử trung tâm

Có nhiều phương pháp phân cụm khác nhau như: phương pháp Ward, ISODATA v.v [11,12,14,15,50]. Ở đây luận án áp dụng thuật toán phân cụm

ISODATA [2] cho tập các điểm ứng cử trung tâm dựa trên một ngưỡng phân cụm xác định, tức là phân cụm tập các điểm trung tâm thành một số cụm sao cho khoảng cách lớn nhất trong mỗi cụm nhỏ hơn một ngưỡng xác định, ngưỡng phân cụm này được xác định theo thực nghiệm tùy theo từng loại chất liệu. Quá trình phân cụm được dừng lại khi khoảng cách giữa các cụm đạt đến một ngưỡng xác định. Sau khi phân cụm ta thu được một số cụm và đếm số các điểm trong mỗi cụm, nếu số phần tử trên mỗi cụm nhiều hơn một ngưỡng bầu cử theo thực nghiệm thì hệ thống kết luận mẫu chất liệu  $\mathcal{M}$  có trên ảnh vào  $\mathcal{I}$ .

Mẫu chất liệu và bản thân chất liệu trong ảnh có thể xuất hiện nhiều vùng, có thể là các vùng liên tục, có thể là các vùng rời rạc. Vì vậy, để xác định tất cả các vị trí của mẫu chất liệu trong ảnh thì phải thực hiện lặp lại nhiều lần tìm kiếm mẫu chất liệu trong ảnh. Mỗi lần tìm thấy mẫu chất liệu xuất hiện tại vùng nào đó trong ảnh, loại bỏ vùng đó và tiếp tục tìm kiếm cho các vùng khác trên ảnh.

#### 2.3.4. Thuật toán phát hiện mẫu chất liệu DMBLIF

- **Vào:** Mẫu chất liệu  $\mathcal{M}$  (dưới dạng một bức ảnh), ảnh cần tìm chất liệu  $\mathcal{I}$ .
- **Ra:** Tập  $R$  gồm các vùng chứa mẫu chất liệu  $\mathcal{M}$  trên ảnh  $\mathcal{I}$ .
- Các bước chính của thuật toán:

**Bước 1:** Khởi tạo các tham số ngưỡng phân cụm và ngưỡng phân lớp  $\varepsilon, \gamma$ . Chọn điểm trung tâm  $M_c$  trên mẫu chất liệu  $\mathcal{M}$ . Khởi tạo  $R := \{\emptyset\}$ , khởi tạo tập các điểm trung tâm ứng cử trên ảnh  $\mathcal{I}$  là  $I_c := \{\emptyset\}$ .

**Bước 2:** Trích chọn đặc trưng bất biến địa phương cho mẫu chất liệu  $\mathcal{M}$  và ảnh  $\mathcal{I}$  như phần 2.2, giả sử ta được tập  $FM = \{F_M^i\}_{i=1}^K$  và  $FI = \{F_I^j\}_{j=1}^L$  tương ứng.

**Bước 3:** Với mỗi đặc trưng  $F_M^i \in FM, i = 1 \dots K$ , thực hiện:

**Bước 3.1:** Tìm đặc trưng  $F_I^{j'} \in FI$  giống nhất với  $F_M^i$  theo công thức (2.7).

**Bước 3.2:** Tính các thành phần của véc tơ định vị  $\Delta_i = (\Delta_x, \Delta_y)$  dựa vào  $F_M^i$  và  $M_c$  (xem hình 2.3).

**Bước 3.3:** Dựa vào  $F_i^{j'}$  đã tìm được ở bước 3.1 và  $\Delta_i$  đã tìm được ở bước 3.2, tính điểm trung tâm ứng cử  $I_c^i = (X_c^i, Y_c^i)$  trên  $\mathcal{J}$  theo công thức (2.8) và lưu vào tập  $I_c$ , do vậy  $I_c := I_c \cup \{I_c^i\}$ .

**Bước 4:** Phân cụm tập điểm trung tâm ứng cử  $I_c = \{I_c^i\}_{i=1}^K$  (tính được ở bước 3) thành một số cụm dựa vào thuật toán  $ISODATA(I_c, \varepsilon)$  với  $\varepsilon$  là ngưỡng khoảng cách lớn nhất trong mỗi cụm.

**Bước 5:** Với mỗi cụm tìm được trong bước 4, kiểm tra nếu số điểm lớn hơn một ngưỡng phân lớp  $\gamma$  xác định thì kết luận mẫu chất liệu có trên  $\mathcal{J}$ , và thực hiện:

**Bước 5.1:** Lưu  $Rec(FC)$  (vùng bao chứa tập các điểm đặc trưng  $FC$  tương ứng của cụm điểm trung tâm ứng cử được chọn trong ảnh  $\mathcal{J}$ ),  $R := R \cup \{Rec(FC)\}$ .

**Bước 5.2:** Loại bỏ các đặc trưng tương ứng của cụm trong tập  $FI$ , ( $FI = FI \setminus FC$ )

**Bước 5.3:** Kiểm tra nếu  $FI \neq \emptyset$  thì quay lại bước 3. Ngược lại thì kết thúc.

- Nhận xét:
  - Thuật toán DMBLIF luôn đảm bảo tính dừng: Nếu toàn bộ ảnh  $\mathcal{J}$  chứa chất liệu  $\mathcal{M}$  thì theo bước 5.2 số các đặc trưng trong tập  $FI$  sẽ giảm dần về 0 và thuật toán sẽ dừng theo bước 5.3. Nếu ảnh  $\mathcal{J}$  không hoàn toàn chứa chất liệu  $\mathcal{M}$  thì luôn tồn tại một thời điểm (không còn vùng nào của ảnh  $\mathcal{J}$  chứa mẫu chất liệu  $\mathcal{M}$ ) thì điều kiện của bước 5 không thỏa mãn, nghĩa là thuật toán sẽ dừng ở bước 5. Vì vậy, thuật toán luôn đảm bảo tính dừng.
  - Thuật toán DMBLIF luôn đảm bảo tính đúng đắn: Ở bước 2, chúng ta luôn trích chọn được đặc trưng cho các tập  $FM$  và  $FI$ . Việc tính tập các điểm ứng cử trung tâm  $I_c$  từ bước 3.1 đến bước 3.3 là hoàn toàn chính xác do các công thức (2.7) và (2.8) là các công thức rõ. Ở bước 4, ta

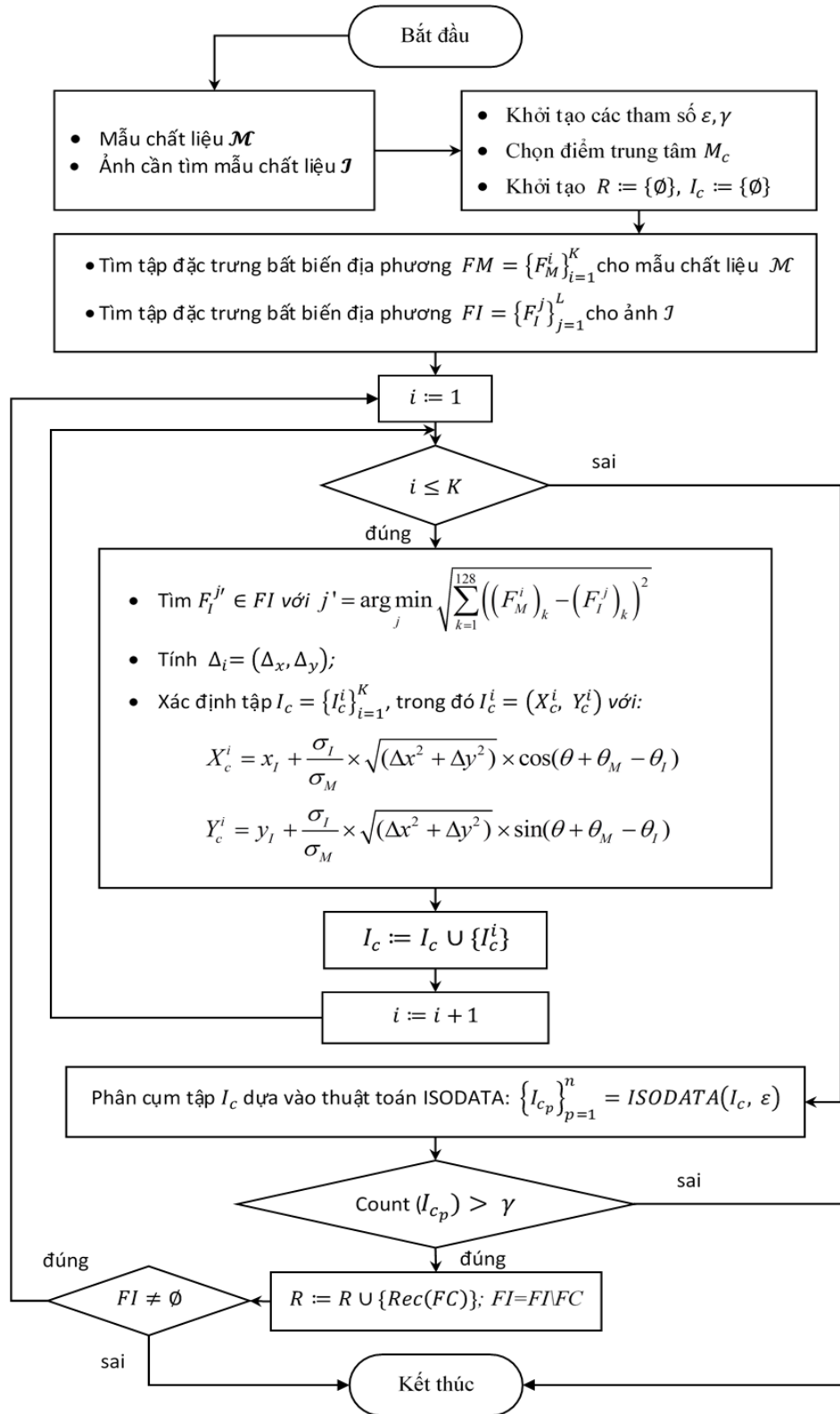
luôn phân cụm được tập  $I_c$  dựa vào thuật toán ISODATA với ngưỡng  $\varepsilon$  luôn đạt được. Nếu mẫu chất liệu  $\mathcal{M}$  có trên ảnh  $\mathcal{J}$  thì việc kiểm tra điều kiện ở bước 5 là đúng, do đó  $Rec(FC) \neq 0$  nghĩa là vùng  $FC$  được xác định là mẫu chất liệu. Nếu mẫu chất liệu  $\mathcal{M}$  không có trên ảnh  $\mathcal{J}$  thì điều kiện ở bước 5 là sai và  $Rec(FC) = 0$  nghĩa là vùng  $FC$  không được xác định là mẫu chất liệu. Do vậy, thuật toán luôn cho kết quả đúng đắn.

- Độ phức tạp tính toán thuật toán DMBLIF:

Độ phức tạp của thuật toán DMBLIF phụ thuộc vào số đặc trưng trích chọn được từ ảnh vào. Giả sử  $K$  là số đặc trưng trích chọn được của mẫu chất liệu và  $L$  là số đặc trưng trích chọn được từ ảnh vào, số phép toán tối đa trong từng bước được tính như sau:

Trong một bước tìm chất liệu ta có độ phức tạp là  $O(K.L)$ . Số lần lặp để tìm chất liệu là  $O(\frac{L}{K})$ . Do đó, độ phức tạp trong trường hợp xấu nhất của cả quá trình là:  $O((\frac{L}{K})(K.L)) \approx O(L^2)$ .

- Thuật toán được biểu diễn dưới dạng sơ đồ khối:



### 2.3.5. Thực nghiệm

- Cơ sở dữ liệu thử nghiệm:

Chúng tôi lấy cơ sở dữ liệu từ video của MIT – CSAIL (MIT Traffic Database) của khoa Điện tử - Trường Đại học Hồng Kông, được xây dựng nhằm phục vụ cho việc phân tích hoạt động đám đông trên đường phố. Đoạn video này dài hơn 90 phút được chia thành 20 clip, mỗi clip dài 4 phút 36 giây, với độ phân giải  $720 \times 480$ , được thu từ camera tĩnh đặt trên cao với góc nghiêng khoảng  $30^\circ$  thu cảnh một góc phố trong điều kiện ánh sáng ban ngày, khung nền tĩnh, các hoạt động chủ yếu là người đi bộ và xe cộ qua lại. Trong video này chúng tôi lấy ra 2014 ảnh có mẫu chất liệu cỏ và tán cây của 3 clip liên tiếp, các ảnh lấy ra cách nhau 10 frame. Đồng thời chúng tôi cũng tự thu nhận được 1470 ảnh có mẫu chất liệu gỗ, 693 ảnh có mẫu chất liệu da và 818 ảnh có mẫu chất liệu tóc từ cảnh trong một ngôi nhà cùng với độ phân giải  $720 \times 480$  và điều kiện ánh sáng thay đổi lớn, tỉ lệ thay đổi từ 0.8 đến 1.2, góc quay thay đổi từ 0 đến 90 độ một cách ngẫu nhiên.

- Cách đánh giá:

Kỹ thuật đề xuất được cài đặt bằng công cụ lập trình Visual C++ 2008 (*Chi tiết xem phụ lục AI*), sử dụng thư viện mã nguồn mở Open CV và sử dụng tập dữ liệu thử nghiệm MIT-CSAIL của Khoa Điện tử - Trường Đại học Hồng Kông và trên tập dữ liệu chúng tôi thu nhận được. Chúng tôi thử nghiệm trên mẫu chất liệu cỏ, da, tóc và gỗ được cắt ra từ các tập loại ảnh trên. Chúng tôi đánh giá tính đúng, sai bằng mắt thường quan sát, nếu phát hiện được số vùng chứa mẫu chất liệu lớn hơn ngưỡng nào đó theo tỉ lệ xác định của số vùng chất liệu phát hiện được so với tổng số vùng chứa mẫu chất liệu trên ảnh thì được cho là đúng, ngoài ra được cho là sai.

- Chương trình gồm các pha chính:
  - Đọc ảnh mẫu chất liệu
  - Đọc ảnh vào cần tìm mẫu chất liệu
  - Trích chọn đặc trưng bất biến của mẫu chất liệu và ảnh vào

- Với mỗi đặc trưng của mẫu chất liệu, tìm đặc trưng giống với nó nhất ở trong ảnh vào.
- Kiểm chứng các cặp đặc trưng đã đối sánh được trong mẫu chất liệu và trong ảnh vào bằng cấu trúc tương quan hình học được trình bày trong thuật toán.
- Phân cụm tập các điểm ứng cử trung tâm trên ảnh trong ảnh vào theo thuật toán ISODATA với khoảng cách cực đại  $\varepsilon = 2$ .
- Xác định mẫu chất liệu có trên ảnh vào dựa vào số đặc trưng trên mỗi cụm theo ngưỡng  $\gamma = 3$ .

Thử nghiệm trên các tập ảnh đã chọn, chúng tôi nhận được kết quả như sau:

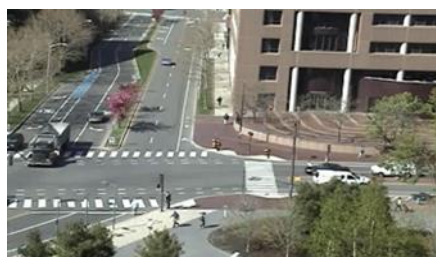
**Bảng 2.1. Kết quả đánh giá thuật toán DMBLIF**

Mẫu chất liệu thử nghiệm	Kết quả phát hiện mẫu chất liệu bằng thuật toán DMBLIF				
	Số ảnh	Đúng	Sai	Tỉ lệ sai	Độ chính xác
Gỗ	1470	1446	24	1,60	98,40%
Cỏ	2014	1994	20	1,00	99,00%
Da	693	432	261	37,66	62,33%
Tóc	818	696	122	14,92	85,08%

Quá trình phát hiện được sử dụng một ngưỡng thực nghiệm, tùy từng loại mẫu chất liệu cụ thể sẽ lấy các ngưỡng phân cụm và ngưỡng phân loại khác nhau. Ở thực nghiệm này, với mẫu chất liệu gồm một ảnh chất liệu, ảnh vào có kích thước  $720 \times 480$ , thời gian thực hiện phát hiện mẫu chất liệu trong mỗi ảnh vào đều là thời gian thực, thời gian cụ thể để phát hiện các chất liệu khác nhau sẽ khác nhau vì chúng còn phụ thuộc vào số các điểm đặc trưng địa phương được trích chọn tùy theo từng mẫu chất liệu cụ thể sẽ có các ngưỡng khác nhau nhưng tất cả đều thực hiện được trên thời

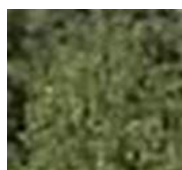


gian thực. Thực nghiệm cho thấy với mẫu chất liệu cỏ cây và gỗ cho độ chính xác cao hơn mẫu chất liệu da và tóc, điều này cũng chứng tỏ được ưu điểm của kỹ thuật phát hiện tốt trong trường hợp mẫu chất liệu có cấu trúc hình học cao. Dưới đây là một số hình ảnh minh họa kết quả phát hiện mẫu chất liệu trong các trường hợp mẫu chất liệu bị thay đổi quay và thay đổi tỉ lệ:

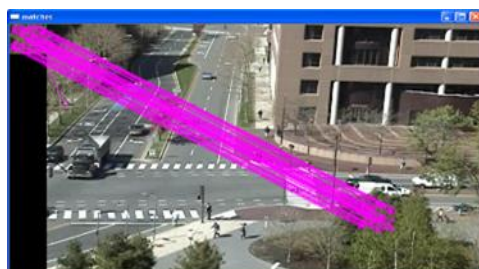


(a)

(b)



(c)



(d)



(e)



(f)



(g)



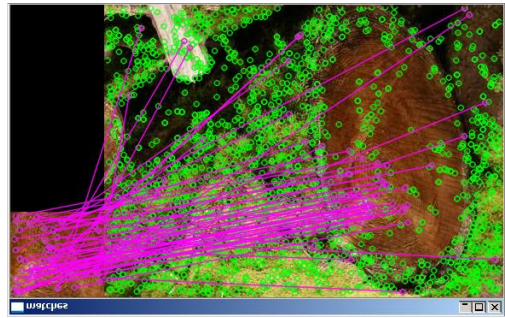
Hình 2.4. Kết quả phát hiện mẫu chất liệu cỏ bằng thuật toán DMBLIF  
 (a) Ảnh vào; (b) Mẫu chất liệu cỏ ; (c) Kết quả phát hiện mẫu (b); (d) Mẫu cỏ bị quay; (e) Kết quả phát hiện mẫu (d); (f) Mẫu cỏ thay đổi quay và tỉ lệ; (g) Kết quả phát hiện mẫu (f).



(a)



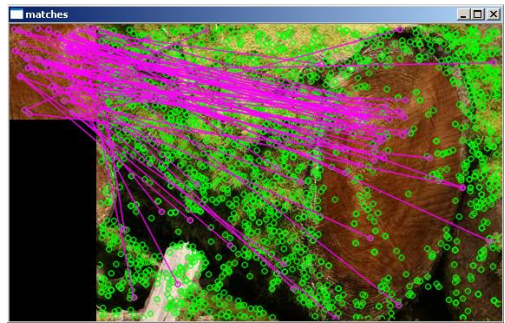
(b)



(c)



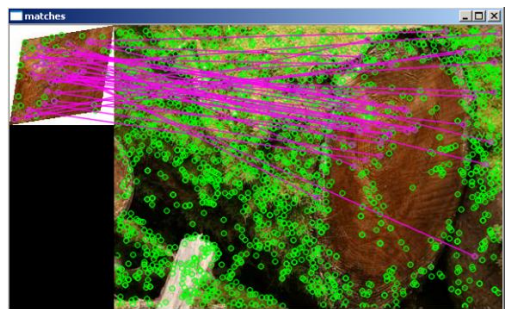
(d)



(e)



(f)



(g)

Hình 2.5. Kết quả phát hiện mẫu chất liệu gỗ bằng thuật toán DMBLIF

(a) Ảnh vào; (b) Mẫu gỗ bị thay đổi tỉ lệ; (c) Kết quả phát hiện khi mẫu gỗ thay đổi quay và tỉ lệ ; (d) Mẫu gỗ bị quay ; (d) Kết quả phát hiện mẫu gỗ đã bị thay đổi quay ; (e) Mẫu gỗ bị thay đổi tỉ lệ và quay ; (f) Kết quả phát hiện mẫu gỗ bị thay đổi quay và thay đổi tỉ lệ.

Phát hiện mẫu chất liệu dựa trên các đặc trưng bất biến địa phương là một cách tiếp cận đầy hứa hẹn. Vì mẫu chất liệu chính là một phần của đối tượng được biểu diễn trong ảnh nên cũng bị thay đổi bởi một số phép biến đổi ảnh. Điểm khó khăn nhất đó là chất liệu ảnh bị thay đổi bởi tỉ lệ, nhưng phương pháp này đã giải quyết được. Nhược điểm của cách tiếp cận này là không phát hiện được cho các mẫu chất liệu chỉ đồng nhất trên từng pixel ví dụ như chất liệu da qua thực nghiệm độ chính xác chỉ dưới 70%. Nếu thực nghiệm trên các mẫu chất liệu có cấu trúc hình học bề mặt cao và có chất lượng tốt, hay các ảnh nhân tạo với độ sáng không thay đổi nhiều thì phương pháp này cho kết quả đạt đến gần 100%.

## **2.4. Phát hiện ảnh số giả mạo dựa vào thuật toán DMBLIF**

Ảnh số ngày càng dễ sửa chữa và hiệu chỉnh do sức mạnh của các phần mềm soạn thảo và xử lý ảnh. Ngày nay, các chương trình xử lý ảnh có thể thêm vào hoặc bỏ đi các đặc trưng của ảnh mà không để lại các dấu hiệu về sự giả mạo. Điều đó có nghĩa là việc phát hiện ảnh số giả mạo là vấn đề đặt ra cấp bách và càng trở nên khó khăn [62,65,71].

### **2.4.1. Ảnh số giả mạo và các dạng ảnh số giả mạo cơ bản**

Ngày nay, ảnh số là phương tiện truyền thông được sử dụng phổ biến và rộng rãi, nó chứa đựng nhiều thông tin, cũng là cách để biểu diễn thế giới xung quanh ta một cách dễ dàng và cô đọng. Một câu hỏi đặt ra là ảnh này có độ tin cậy bao nhiêu? Rõ ràng, thông tin hình ảnh có tác động quan trọng đến xã hội chúng ta, đóng vai trò cốt yếu trong đời sống con người. Việc xác thực ảnh là bài toán quan trọng trong nhiều lĩnh vực xã hội, chẳng hạn trong phòng xử án, tính thật của một bức ảnh đóng vai trò cốt yếu vì nó chính là một bằng chứng; hàng ngày các bài báo hay tạp chí đều chứa nhiều hình ảnh, trong vật lý và y học thì có nhiều quyết định đều phải dựa vào các ảnh số [41,57,65].

Hơn nữa, hiện nay do sự tiến bộ của các máy tính hiệu năng cao, giá thành thấp, giao diện người và máy thân thiện hơn, cũng như các phần mềm điều khiển và

thao tác trên ảnh số là khá mạnh và đơn giản. Sức mạnh của các chương trình xử lý ảnh số như PhotoShop giúp việc tạo ra các ảnh giả mạo từ một hay nhiều ảnh khác nhau trở nên dễ dàng [71,73]. Điều đó, đồng nghĩa với việc phát hiện ảnh giả mạo là bài toán khó khăn và ngày càng trở nên cấp thiết hơn.

#### 2.4.1.1. Ảnh số giả mạo

Ảnh giả mạo được xem là ảnh không có thật, việc có được ảnh là do sự nguy tạo bởi các chương trình xử lý ảnh hoặc quá trình thu nhận. Một trong những ví dụ cho việc giả mạo này là Hình 2.6. Hình ảnh được tạo lập từ 3 bức ảnh: Nhà trắng, Bill Clinton và Saddam Hussein. Hình ảnh Bill Clinton và Saddam Hussein được cắt và dán vào bức ảnh Nhà trắng, các hiệu ứng về bóng và ánh sáng cũng được tạo ra làm cho bức ảnh có vẻ nhìn hoàn toàn như thật.



Hình 2.6. Minh họa về việc giả mạo ảnh

Giả mạo ảnh nhằm vào nhiều mục đích trong đó có việc vu cáo, tạo ra các tin giật gân, đánh lừa đối thủ, làm sai lệch chứng cứ phạm tội v.v. Việc phát hiện và

chống giả mạo ảnh số là một chủ đề ngày càng được quan tâm bởi các nhóm nghiên cứu trên thế giới như: TS. Chung-Sheng Li ở IBM T. J. Watson Research Center, TS. Jessica Fridrich ở Department of Electrical and Computer Engineering, SUNY Binghamton, Binghamton, NY 13902-6000 hay TS. David Rosenthal v.v.

Trong việc chống giả mạo ảnh, người ta đã nghiên cứu các kỹ thuật về tạo bản quyền ảnh trên cơ sở giấu các thông tin cần thiết vào bức ảnh trước khi phát hành để tránh tình trạng sao chép bất hợp pháp hoặc để tiện cho việc phát hiện sửa đổi hoặc cắt ghép sau này. Theo cách tiếp cận này, các thiết bị máy ảnh số và camera ngày nay thường bổ sung các tính năng bản quyền ngay trong quá trình thu nhận ảnh. Dựa vào đó, sau này có thể biết được nguồn gốc của bức ảnh. Nhờ đó mà có thể biết được đó có phải là ảnh gốc hay không. Phát hiện ảnh giả mạo là một vấn đề khó hơn, ảnh giả mạo thường chia làm hai loại chính. Ảnh giả nhưng thật, ở đây hiện trường được dựng thật và sau đó việc thu nhận ảnh là thật. Loại thứ hai là ảnh giả được tạo lập trên cơ sở các phần của ảnh gốc thật như hình 2.6 hoặc được cắt dán để thêm vào hay che đi các chi tiết trên ảnh. Trong khuôn khổ ứng dụng Công nghệ thông tin, phát hiện ảnh số giả mạo chỉ quan tâm loại ảnh số giả mạo loại thứ hai. Loại ảnh số giả mạo thứ hai lại được phân thành nhiều loại cơ bản tương ứng với các cách làm giả mạo khác nhau.

#### **2.4.1.2. Các dạng ảnh số giả mạo cơ bản**

Trong dạng ảnh giả mạo loại thứ hai có thể chia ra làm các loại: Ghép ảnh, che phủ đối tượng, bổ sung thêm đối tượng. Hình 2.7a minh họa ảnh ghép từ 2 ảnh riêng rẽ với cùng tỉ lệ. Rõ ràng là nếu chỉ ra đây là ảnh thật hay ảnh giả mạo thì cũng có nghĩa là chứng minh được mối quan hệ giữa họ. Hình 2.7b là ảnh ghép từ hai ảnh có sự thay đổi tỉ lệ. Hình ảnh này nếu không chứng minh là giả tạo thì sẽ phải có cách nhìn khác về tiến hóa của loài người?



(a)



(b)

*Hình 2.7. Ghép ảnh từ 2 ảnh riêng rẽ*

*(a) Ảnh ghép từ hai ảnh riêng rẽ; (b) Ảnh ghép từ hai ảnh có thay đổi tỉ lệ*

Một dạng khác thường thấy nữa của ảnh giả mạo loại hai là việc bớt đi các đối tượng trong ảnh, việc bớt đi các đối tượng trong ảnh có thể được xem như là che phủ đối tượng hoặc xóa đi đối tượng. Hình 2.8a là ảnh gốc có 2 chiếc ô tô, một ô tô con và một ô tô tải. Hình 2.8b là ảnh 2.8a bị giả tạo với việc che phủ chiếc ô tô tải bởi một cành cây cũng lấy từ chính trong ảnh. Hình 2.8c là ảnh gốc với chiếc trục thẳng nhỏ còn hình 2.8d chính là ảnh gốc 2.8c đã được bỏ đi đối tượng trục thẳng. Cả hai trường hợp giả mạo này đều được thực hiện từ một ảnh nên độ tương đồng về ánh sáng và bóng là như nhau, do đó bằng mắt thường rất khó xác định.



(a)



(b)



(c)



(d)

Hình 2.8. Ảnh che phủ và bỏ đi đối tượng  
(a), (c) Ảnh gốc; (b), (d) Ảnh đã che phủ đối tượng;



(a)



(b)

Hình 2.9. Ảnh bổ sung đối tượng  
(a) Ảnh gốc; (b) Ảnh bổ sung đối tượng

Hình 2.9 thể hiện một dạng khác nữa thường thấy của giả mạo ảnh, đó là việc bổ sung thêm đối tượng. Hình 2.9a là ảnh gốc chỉ có một chiếc máy bay trực thăng, nhưng trong hình 2.9b đã được bổ sung thêm thành 3 chiếc trực thăng ở các vị trí khác nhau. Các trực thăng này chính là được copy từ trực thăng gốc nên góc độ và hướng là giống nhau, do đó khó cho việc xác định.

#### 2.4.2. Thuật toán phát hiện ảnh số giả mạo KPFIimage

Phát hiện ảnh giả mạo là kiểm tra xem một bức ảnh cho trước có phải bị làm giả hay không. Một ảnh có thể nghi ngờ là giả mạo nếu trong bức ảnh đó có các phần

chính là bản sao của một phần nào đó trong ảnh. Các cách tiếp cận tìm các bản sao của một phần con trong ảnh thường dựa trên việc so sánh các đặc trưng trích chọn được từ ảnh như cường độ ảnh, đặc trưng đường biên (contour), đáp ứng của các bộ lọc hay các đặc trưng địa phương ảnh. Xuất phát từ thuật toán Exact Match của TS. Jessica Fridich [40] đã phát hiện được các phần làm giả có cùng kích thước và tỉ lệ. Song các phần mềm xử lý ảnh hiện nay có thể tạo ra các vùng giả mạo bằng cách thay đổi kích thước tỉ lệ một cách rất dễ dàng để sự ngụy tạo khó phát hiện được, trường hợp này thì thuật toán Exact Match không giải quyết được. Để khắc phục được vấn đề này một cải tiến của Exact Match đó là Exact Match\* [28] của tác giả luận án đề xuất, thuật toán này thực hiện việc đối sánh các khối bao nhằm dò tìm các bản sao xuất hiện trong ảnh trên cơ sở tìm các hệ số cơ giãn thông qua các kỹ thuật nội suy mà các phần mềm xử lý ảnh chuyên dùng, nhưng với trường hợp các bản sao bị thay đổi bởi phép quay thì không thực hiện được. Trên cơ sở nghiên cứu về việc tìm kiếm mẫu chất liệu trong ảnh, các mẫu chất liệu này có thể bị thay đổi bởi tỉ lệ, ánh sáng hay các phép biến đổi hình học như phép quay, dịch chuyển v.v. Do đó nếu biết một vùng nào đó trong ảnh là nghi ngờ giả mạo thì có thể xem nó như là một mẫu chất liệu và tiến hành tìm kiếm các bản sao của mẫu chất liệu này trong ảnh. Mặc khác nhằm giải quyết bài toán phát hiện ảnh giả mạo trong cả trường hợp bản sao giả mạo bị thay đổi bởi phép quay, luận án đã đề xuất ứng dụng phát hiện mẫu chất liệu trong ảnh cho bài toán phát hiện ảnh giả mạo dạng cắt dán là dạng ảnh giả được tạo ra bằng cách thêm hoặc bớt đi một hoặc một số vùng trong ảnh và đã đề xuất áp dụng kỹ thuật phát hiện mẫu chất liệu dựa vào đặc trưng bất biến địa phương nhằm tìm ra các bản sao của các vùng được xem là mẫu chất liệu nghi ngờ giả mạo. Thuật toán được đề xuất trên cơ sở này đã giải quyết được hầu hết các trường hợp làm giả như thay đổi kích thước, tỉ lệ và các phép biến đổi hình học trong đó có phép quay.

Đã có nhiều nhóm nghiên cứu trong việc phát hiện ảnh giả mạo dạng cắt dán. Một trong những nhóm tiêu biểu là J. Fridich [40] đưa ra thuật toán "Exact Match" một thuật toán phát hiện ảnh giả mạo dựa vào đối sánh khối bao. Thuật toán này nhằm



mục đích phát hiện ảnh giả mạo trong trường hợp ảnh bị cắt dán bởi các phần khác trong ảnh. Tuy nhiên, nó lại không giải quyết được trường hợp các đối tượng giả mạo sau khi được trích ra từ một phần nào đó của ảnh đã được thay đổi kích thước hoặc quay đi một góc cho phù hợp, mà đây là cách thường làm ra ảnh giả. Hình 2.10a là ảnh gốc với một chiếc máy bay trực thăng. Hình 2.10b được tạo ra từ ảnh 2.10a bằng cách bổ sung thêm thành bốn chiếc trực thăng ở các vị trí khác nhau. Các trực thăng này chính là được sao chép từ trực thăng gốc.



(a)



(b)

Hình 2.10. Ảnh giả mạo cắt dán bởi bổ sung đối tượng  
(a) Ảnh gốc; (b) Ảnh giả mạo bổ sung đối tượng

Trong [28] đã đề xuất một cải tiến cho thuật toán Exact match trên cơ sở nghiên cứu các kỹ thuật nội suy hiện đang được PhotoShop sử dụng. Thuật toán này có khả năng phát hiện đối với các ảnh giả mạo dạng cắt dán từ chính một ảnh và có sự thay đổi về kích thước đối với các đối tượng bị cắt dán. Myna và các cộng sự [73] đã đưa ra phương pháp phát hiện ảnh giả mạo dựa trên phép biến đổi Wavelet, nhưng vẫn không giải quyết được trong trường hợp đối tượng được dán đã bị phóng to hay thu nhỏ lại.

Ở đây, luận án đề xuất kỹ thuật phát hiện ảnh giả mạo dạng cắt dán dựa trên cơ sở phát hiện mẫu chất liệu trong ảnh dựa vào đặc trưng bất biến địa phương để phát hiện ảnh giả mạo dạng cắt dán trong trường hợp vùng nghi ngờ giả mạo bị thay đổi tỉ lệ, thay đổi kích thước và còn cả sự thay đổi về sự dịch chuyển và quay.

Ảnh số giả mạo dạng copy/move thuộc dạng ảnh giả mạo được tạo ra bằng cách cắt dán từ các ảnh khác nhau hoặc trên cùng một ảnh. Dạng ảnh giả này, sử dụng một phần của ảnh được copy và dán vào một phần khác trên cùng một ảnh. Điều này thực hiện nhằm mục đích làm cho đối tượng nào đó bị biến mất, bằng cách phủ lên đó phần được copy từ chỗ khác. Các vùng kết cấu như cỏ, tán lá, sỏi đá hoặc vải với các mẫu không đều đặn là các vùng được dùng để sao chép và dán lên đối tượng sẽ biến mất vì nó giống như pha trộn với nền, nên mắt thường không nhìn thấy được sự nghi ngờ nào. Do các phần được sao chép và dán trên cùng một ảnh nên các thành phần nhiễu, tông màu và hầu hết các thuộc tính quan trọng khác sẽ tương đồng với phần còn lại vì vậy sẽ rất khó phát hiện sự không tương thích bằng các phương pháp thống kê. Để làm cho ảnh giả thật hơn, người ta có thể phóng to, thu nhỏ hay quay đi một góc trước khi dán.

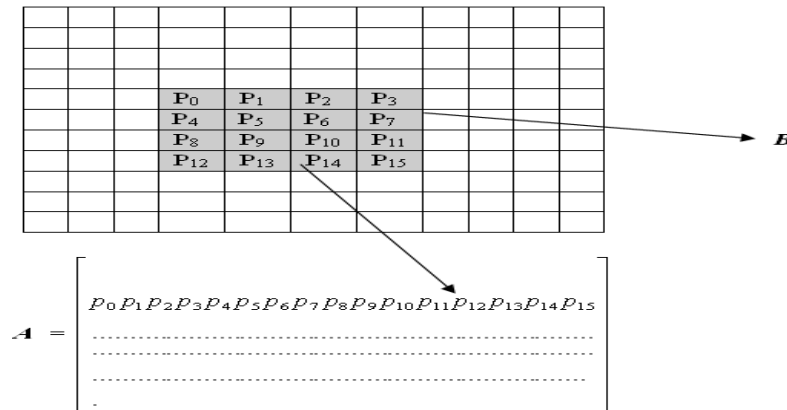
#### **2.4.2.1. Thuật toán phát hiện ảnh số giả mạo Exact Match**

Thuật toán được đề xuất bởi nhóm J. Fridich [40]. Ý tưởng của thuật toán là sử dụng các khối bao để đối sánh nhằm tìm ra các vùng giống nhau trên toàn bộ bức ảnh.

Giả sử bức ảnh có kích thước  $M \times N$ , với  $B$  là kích thước nhỏ nhất của khối bao mà người dùng định nghĩa để đối sánh. Với mỗi điểm ảnh ta xác định được một khối bao ma trận  $B \times B$  điểm ảnh. Như vậy, với bức ảnh  $\mathcal{J}(M \times N)$  ta xác định được  $(M - B + 1) \times (N - B + 1)$  khối bao. Với mỗi khối bao, ta lưu các phần tử thuộc khối bao vào một hàng của một ma trận  $A$ . Duyệt trên toàn bộ bức ảnh, ta sẽ được một ma trận  $A$  với  $(M - B + 1) \times (N - B + 1)$  hàng và  $B \times B$  cột.

Hai hàng giống nhau trong ma trận  $A$  tương đương với 2 khối bao giống nhau trong ảnh. Ta sắp xếp các hàng trong ma trận  $A$  theo thứ tự từ điển. Yêu cầu này sẽ được thực hiện trên  $M \times N \times \log_2(M \times N)$  bước. Sau đó, tìm kiếm bằng cách duyệt  $M \times N$  hàng của ma trận đã qua sắp xếp  $A$  và tìm kiếm hai hàng giống nhau liên tiếp.

Kết quả thuật toán sẽ tìm kiếm được tập các vùng bao giống nhau, minh chứng cho việc ảnh đã bị cắt dán.



Hình 2.11. Tìm kiếm khối bao của thuật toán Exact Match

Kết quả trong hình 2.12 minh họa cho thuật toán này.

Hình 2.12a là ảnh gốc với một chiếc trực thăng. Hình 2.12b là ảnh giả được tạo từ ảnh gốc và thêm vào 3 chiếc trực thăng, trong đó có 1 chiếc được copy mà không thay đổi kích thước 1 chiếc có thay đổi kích thước, 1 chiếc bị quay và 1 chiếc bị thay đổi tỉ lệ và quay. Hình 2.12c là ảnh kết quả phát hiện giả mạo nhờ thuật toán Exact match, kết quả chỉ ra những chiếc trực thăng giả mạo không thay đổi kích thước bị phát hiện.



(a)

(b)

(c)

Hình 2.12. Kết quả phát hiện giả mạo bằng thuật toán Exact match

(a) Ảnh gốc; (b) Ảnh giả mạo cắt dán;

(c) Các vùng giả mạo được phát hiện bởi thuật toán Exact Match

### 2.4.2.2. Thuật toán KPFIImage

Để phát hiện các vùng ảnh giả mạo, ta phải tìm ra các cặp vùng ảnh tương quan (cặp các vùng ảnh “giống nhau”) với những tỉ lệ khác nhau thông qua hệ số tỉ lệ  $\alpha$  và các phép biến đổi nội suy  $f$  [48,49]. Thuật toán được đề xuất trong luận án này là sự phát triển và mở rộng của hai thuật toán Exact Match và thuật toán Exact Math\*[28]. Vùng ảnh thứ nhất được xác định để lấy đối sánh, sẽ được thực hiện dựa theo thuật toán Exact Match để chỉ định kích thước khối bao nhỏ nhất  $B \times B$ . Vùng ảnh thứ hai được xác định dựa theo thuật toán Exact Match\*, tức là được xác định dựa trên phép nội suy  $f$  và hệ số tỉ lệ  $\alpha$ . Sự khác biệt trong việc đối sánh giữa hai vùng ảnh so khớp các đặc trưng bất biến được trích chọn như phần 2.2. Vì các đặc trưng này bất biến đối với phép tỉ lệ và phép quay cục bộ, nên ta sẽ tìm được các cặp vùng ảnh tương quan trong cả trường hợp vùng ảnh bị xoay, hay nói khác hơn đối tượng được dán bị thay đổi bởi phép quay ảnh.

Theo [28], cặp vùng  $(\mathcal{R}, \mathcal{R}_\epsilon)$  được gọi là cặp tương quan nếu với mỗi vùng  $\mathcal{R}$ , tồn tại ít nhất một vùng  $\mathcal{R}_\epsilon$  và một ánh xạ  $f$  sao cho:  $f: \mathcal{R} \xrightarrow{\alpha} \mathcal{R}_\epsilon$ ,  $f$  là phép nội suy và  $\alpha$  là hệ số tỉ lệ. Giả sử bức ảnh có kích thước  $M \times N$ ,  $\alpha$  là hệ số tỉ lệ xác định,  $f$  là phép biến đổi nội suy tuyến tính xác định và  $B \times B$  là kích thước khối bao nhỏ nhất có thể được chỉ định, kích thước khối bao  $B$  là do người dùng chỉ định cho phù hợp với từng ảnh khác nhau, có thể xác định qua thực nghiệm để tìm kích thước phù hợp nhất.

Duyệt toàn bộ bức ảnh theo chiều (left-top)  $\rightarrow$  (right-bottom). Với mỗi điểm ảnh ta xác định được hai khối bao tương ứng có kích thước lần lượt là  $B \times B$  và  $\alpha B \times \alpha B$  bao quanh nó. Do vậy, khi duyệt trên toàn bộ bức ảnh sẽ có tất cả  $(M - B + 1) \times (N - B + 1)$  khối bao có kích thước  $B \times B$  và  $(M - \alpha B + 1) \times (N - \alpha B + 1)$  khối bao có kích thước  $\alpha B \times \alpha B$ .

Gọi tập hợp các khối bao có kích thước  $B \times B$  là  $\mathcal{R}_1$  và tập hợp các khối bao có kích thước  $\alpha B \times \alpha B$  là  $\mathcal{R}_2$ , giả sử  $\mathcal{R}_1$  và  $\mathcal{R}_2$  được biểu diễn như sau:

$$\mathfrak{R}_1 = \{a_1, a_2, \dots, a_i, \dots, a_m\}$$

$$\mathfrak{R}_2 = \{b_1, b_2, \dots, b_j, \dots, b_n\}$$

Trong đó  $a_i$  là khối bao thứ  $i$  có kích thước  $B \times B$  với  $m = (M - B + 1) \times (N - B + 1)$ , và  $b_j$  là khối bao thứ  $j$  có kích thước  $\alpha B \times \alpha B$  với  $n = (M - \alpha B + 1) \times (N - \alpha B + 1)$ .

Xét tập hợp  $\mathfrak{R}_1$ , ứng với mỗi phần tử khối bao  $a_i$  có kích thước  $B \times B$  (với  $i = \overline{1, m}$ ), áp dụng kỹ thuật biến đổi nội suy tuyến tính với phép biến đổi  $f$  cho trước và một hệ số tỉ lệ  $\alpha$  đã được định nghĩa ta sẽ được một phần tử khối bao mới  $a'_i$  có kích thước  $\alpha B \times \alpha B$  theo công thức sau:

$$f: a_i \xrightarrow{\alpha} a'_i$$

Như vậy, với tập hợp các khối bao  $\mathfrak{R}_1$  sau khi áp dụng lần lượt phép biến đổi nội suy tuyến tính đối với từng phần tử trong  $\mathfrak{R}_1$  ta sẽ được một tập hợp mới tương ứng là  $\mathfrak{R}'_1 = \{a'_1, a'_2, \dots, a'_i, \dots, a'_m\}$ .

Ta tìm kiếm các cặp khối bao tương quan trong  $\mathfrak{R}_1$  và  $\mathfrak{R}_2$  như sau: Đối sánh lần lượt các khối bao thuộc tập  $\mathfrak{R}'_1$  với các khối bao thuộc tập  $\mathfrak{R}_2$  bằng cách so khớp các đặc trưng bất biến của khối bao  $a'_i$  và  $b_j$ ,  $a'_i = \{F_1^i, F_2^i, \dots, F_u^i\}$  ( $u$  là số đặc trưng của khối bao  $a_i$ ). Nếu tồn tại một cặp khối bao  $(a'_i, b_j)$  với  $a'_i \in \mathfrak{R}'_1$  và  $b_j \in \mathfrak{R}_2$  (trong đó  $i = \overline{1, m}, j = \overline{1, n}$ ) thỏa mãn  $a'_i$  và  $b_j$  là “giống nhau”, ta khẳng định cặp khối bao  $(a_i, b_j)$  là cặp tương quan.

Việc so khớp hai khối bao  $a'_i$  và  $b_j$  được thực hiện dựa trên phương pháp so khớp các đặc trưng bất biến, nghĩa là với mỗi đặc trưng của  $a'_i$  ta tìm đặc trưng của  $b_j$  sao cho khoảng cách giữa  $a'_i$  và  $b_j$  là bé hơn một ngưỡng  $\varepsilon$  cho trước. Lúc này, ta xác định được các cặp đặc trưng giống nhau. Nếu tổng số các cặp đặc trưng giống

nhau đạt một tỉ lệ  $\gamma$  nào đó thì ta nói cặp  $(a'_i, b_j)$  là so khớp được, nghĩa là “giống nhau”.

Vậy tập hợp các cặp khối bao tương quan  $(a_i, b_j)$  trong đó  $a_i \in \mathfrak{R}_1, b_j \in \mathfrak{R}_2$  là các cặp bị nghi là giả mạo. Do mỗi khối bao ứng với một vùng trong ảnh, tập hợp các khối bao tương quan tương ứng với tập hợp các cặp vùng trong ảnh đã được làm giả.

Trên cơ sở đó luận án đưa ra thuật toán tìm ra các khối bao tương quan với nhau nhằm phát hiện các vùng ảnh giả mạo. Thuật toán được trình bày như sau:

**Thuật toán KPFIImage:**

- **Vào:** Ảnh cần kiểm tra giả mạo  $\mathcal{J}$
- **Ra:** Tập  $RF$  gồm tập các vùng nghi ngờ giả mạo trên ảnh  $\mathcal{J}$  nếu có.
- Các bước thực hiện chính của thuật toán:

**Bước 1:** Chọn các thông số đầu vào bao gồm: Kích thước khối bao nhỏ nhất  $B \times B$ , các ngưỡng đánh giá sự “giống nhau”  $\varepsilon$  và  $\gamma$ , các tỉ lệ co giãn theo chiều ngang và chiều dọc là  $p, q$  và hàm nội suy  $f$ . Khởi tạo  $RF := \{\emptyset\}$ .

**Bước 2:** Xác định  $X = \{X^i\}_{i=1}^n$  là tập các ảnh con có kích thước  $B \times B$  của ảnh  $\mathcal{J}$ .

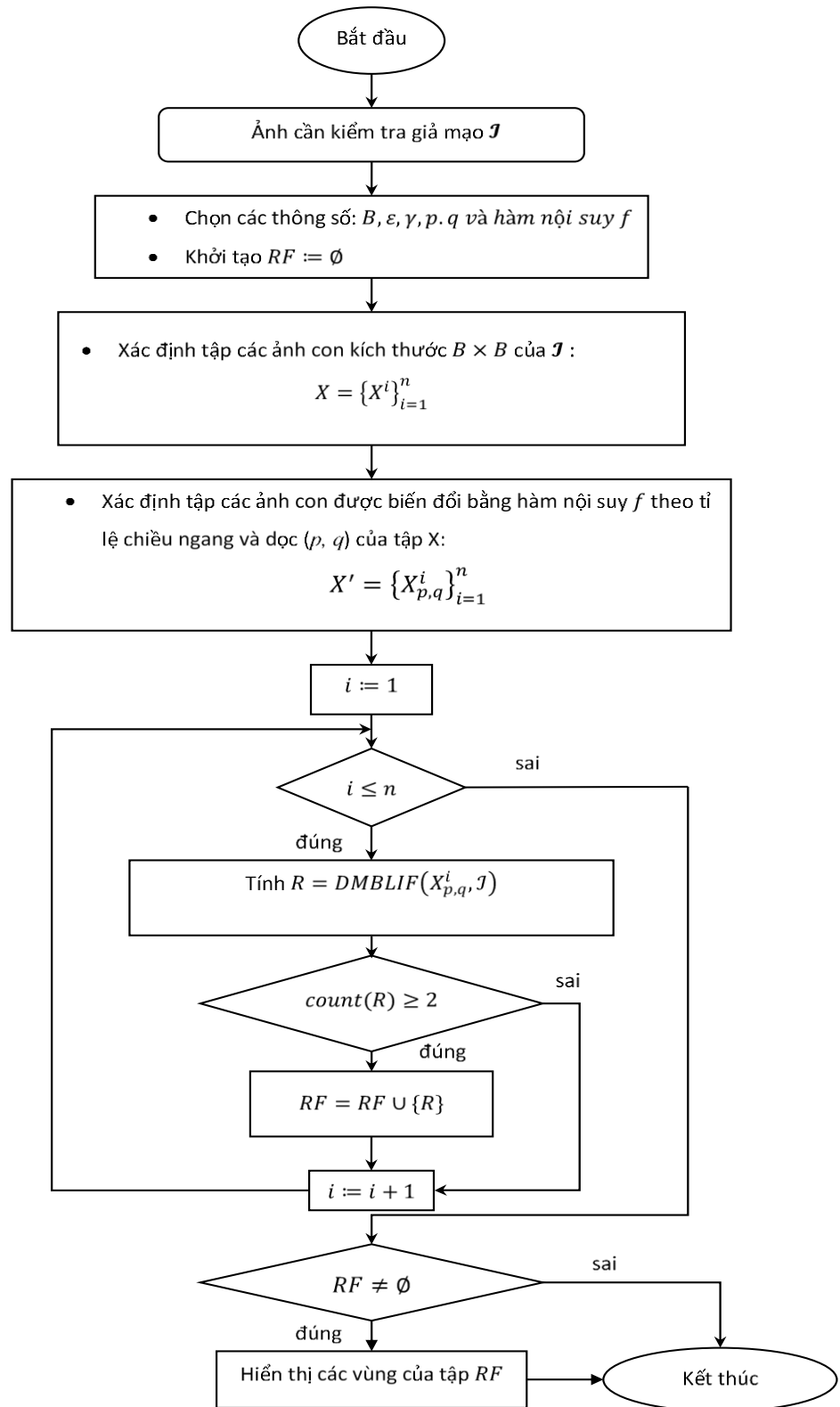
**Bước 3:** Biến đổi mỗi  $X^i \in X$  thành  $X_{p,q}^i \in X'$  theo hàm nội suy  $f$  với các thông số về tỉ lệ co giãn theo chiều ngang và chiều dọc đã chọn trong bước 1.

**Bước 4:** Với mỗi  $X_{p,q}^i \in X'$  ta xem như là một mẫu chất liệu và thực hiện thuật toán  $DMBLIF(X_{p,q}^i, \mathcal{J})$ . Nếu số vùng trả về của  $R = DMBLIF(X_{p,q}^i, \mathcal{J})$  lớn hơn hoặc bằng 2 thì cập nhật các vùng này vào  $RF$ , tức là  $RF = RF \cup \{R\}$ .

**Bước 5:** Nếu  $RF \neq \emptyset$  thì hiển thị các vùng nghi ngờ là giả mạo trên ảnh  $\mathcal{J}$ . Đây là các vùng của tập  $RF$  tìm được sau bước 4.

**Bước 6:** Kết thúc.

- Biểu diễn thuật toán KPFIImage dưới dạng sơ đồ khối:



- Nhận xét
  - Thuật toán KPFIImage luôn đảm bảo tính dừng: Thuật toán  $DMBLIF(X_{p,q}^i, \mathcal{J})$  ở bước 4 luôn đảm bảo tính dừng (đã được chứng minh ở phần trước) nên  $RF$  luôn xác định được. Do đó thuật toán sẽ dừng ở bước 5 và bước 6.
  - Thuật toán KPFIImage luôn đảm bảo tính đúng đắn: Bước 2 và bước 3 của thuật toán hoàn toàn xác định được. Việc áp dụng thuật toán  $DMBLIF(X_{p,q}^i, \mathcal{J})$  ở bước 4 cũng luôn cho kết quả đúng đắn (đã được chứng minh trong phần trước). Nếu  $R \geq 2$  (nghĩa là tồn tại 2 vùng giống nhau hoàn toàn hay nói khác hơn tồn tại 2 vùng đó là nghi ngờ giả mạo) thì  $RF \neq \emptyset$  và do đó các vùng nghi ngờ giả mạo này sẽ được hiển thị trong bước 5, ngoài ra thì không hiển thị gì. Như vậy kết quả của thuật toán hoàn toàn đúng đắn.
- Đánh giá độ phức tạp tính toán của thuật toán KPFIImage

Giả sử ảnh vào kích thước  $N \times N$ , phép duyệt tất cả các khối bao sẽ có độ phức tạp là  $O(N \times N)$ . Với mỗi khối bao thực hiện thuật toán trên ta có độ phức tạp tính toán là  $O(L^2)$  ( $L$  là số đặc trưng địa phương trích chọn được như trong thuật toán DMBLIF). Vậy độ phức tạp của cả quá trình trong trường hợp xấu nhất là  $O(L^2 \times N^2)$ .

### 2.4.3. Thực nghiệm

Luận án tạo một cơ sở dữ liệu gồm 100 ảnh giả dạng copy/move với kích thước khác nhau trong đó có một số ảnh giả được lấy từ bài báo gốc Exact Match, cài đặt thuật toán được đề xuất trên cơ sở dữ liệu ảnh giả này, bằng cách dựa vào kỹ thuật phát hiện mẫu chất liệu trong ảnh dựa vào đặc trưng bất biến địa phương, như đã giới thiệu ở phần 2.1. Kết quả phát hiện được các vùng nghi ngờ giả mạo khi có sự thay đổi ảnh như thay đổi tỉ lệ, quay và thuật toán thực hiện khá nhanh, còn thời gian cụ thể cho từng ảnh phụ thuộc vào kích thước ảnh và tùy từng nội dung ảnh (phụ thuộc



vào số đặc trưng trích chọn được trong ảnh), đồng thời cũng phụ thuộc vào việc chọn lựa kích thước khối bao. Trong thực nghiệm này, luận án đã chọn độ co dãn  $(p_\alpha, q_\alpha) = (2, 2)$ . Hình 2.13, 2.14 và 2.15 là những kết quả minh họa cho thuật toán phát hiện ảnh giả mạo. Hình 2.13a là ảnh 24 bit màu giả mạo được tạo ra bằng cách sinh ra các máy bay bị thay đổi tỉ lệ và quay; Hình 2.13b, c, d là ảnh kết quả phát hiện được thực hiện lần lượt từng lượt từng kỹ thuật. Hình 2.14a là ảnh đa cấp xám giả được tạo ra dựa trên việc cắt dán các con vịt có thay đổi tỉ lệ và quay; Hình 2.14b, c, d là ảnh kết quả phát hiện bằng 3 kỹ thuật khác nhau. Thực nghiệm cho thấy thuật toán không những phát hiện được các vùng giả mạo bị thay đổi tỉ lệ mà còn cho cả các vùng được thay đổi bởi phép quay, thậm chí là thay đổi ánh sáng địa phương. Hình 2.15 là kết quả phát hiện giả mạo theo 3 kỹ thuật Exact Match, Exact Match\* và KPFIImage. Kết quả thực nghiệm cho thấy KPFIImage phát hiện ảnh số nghi ngờ giả mạo được trong trường hợp vùng giả mạo bị xoay và thay đổi tỉ lệ mà hai thuật toán Exact Match và Exact Match\* không thực hiện được, và các vùng giả mạo này trong cả ảnh xám hay ảnh màu đều phát hiện được.



(a)



(b)



(c)



(d)



(e)

Hình 2.13. Kết quả phát hiện ảnh giả mạo cắt dán với ảnh 24 bit màu  
(a) Ảnh gốc; (b) Ảnh giả; (c) Phát hiện giả mạo bằng Exact match;  
(d) Phát hiện giả mạo bằng Exact match\*; (e) Phát hiện giả mạo bằng KPFIimage



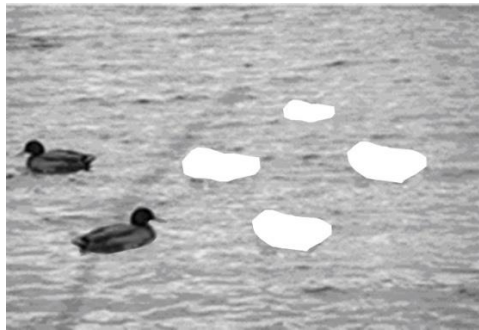
(a)



(b)



(c)



(d)



(e)

Hình 2.14. Kết quả phát hiện ảnh giả mạo cắt dán với ảnh đa mức xám  
(a) Ảnh gốc; (b) Ảnh giả; (c) Phát hiện giả mạo bằng Exact match;  
(d) Phát hiện giả mạo bằng Exact match\*; (e) Phát hiện giả mạo bằng KPFIimage



(a)



(b)



(c)



(d)



(e)

Hình 2.15. Kết quả phát hiện ảnh giả mạo dạng nén  
(a) Ảnh gốc; (b) Ảnh giả với bông hoa bị thay đổi quay;  
(c) Kết quả phát hiện bằng Exact match\* của (b);  
(d) Ảnh giả với bông hoa bị thay đổi tỉ lệ và quay;  
(e) Kết quả phát hiện bằng KPFIimage của (d).

## 2.5. Kết luận chương 2

Chương này đã đề xuất kỹ thuật phát hiện mẫu chất liệu trong ảnh dựa vào đặc trưng bất biến địa phương nhằm giải quyết vấn đề khi hướng thu nhận ảnh thay đổi so với mẫu chất liệu. Kỹ thuật này cho kết quả tốt trong các trường hợp ảnh và mẫu chất liệu bị thay đổi tỉ lệ, quay và ánh sáng đối với các loại chất liệu có độ tương phản cao, không đồng nhất, có tính gồ ghề như: cỏ, gỗ, thảm, vỏ cây v.v. Thực nghiệm trên các tập dữ liệu ảnh có chứa các mẫu chất liệu thông thường cho thấy độ chính xác hơn 95% cho các chất liệu có tính sần sùi như mẫu chất liệu cỏ cây, gỗ. Ở chương này luận án cũng đề xuất ứng dụng thuật toán phát hiện mẫu chất liệu trong ảnh để khắc phục khó khăn về vùng giả mạo bị thay đổi bởi phép quay và phép thay đổi tỉ lệ trong bài toán phát hiện ảnh giả mạo dạng cắt dán theo các định dạng ảnh khác nhau mà một số phương pháp đương thời chưa giải quyết được.

## **Chương 3. PHÁT HIỆN MẪU CHẤT LIỆU DỰA VÀO ĐẶC TRUNG NHIỄU**

Ở chương một đã phân tích, nhiễu chất liệu là một đặc trưng toàn cục và phụ thuộc chủ yếu vào bản thân chất liệu. Biểu diễn và phát hiện mẫu chất liệu dựa vào đặc trưng nhiễu chất liệu ít bị ảnh hưởng bởi sự thay đổi của ánh sáng, đó là một trong những nhân tố ảnh hưởng rất nhiều đến bài toán phát hiện mẫu chất liệu đặc biệt cho những mẫu chất liệu đồng chất như cỏ, đá, giấy v.v. Kỹ thuật phát hiện mẫu chất liệu dựa vào đặc trưng bất biến địa phương phù hợp cho các mẫu chất liệu có tính kết cấu cao (có tính cấu trúc hình học bề mặt cao) và bất biến với một số phép biến đổi hình học nhưng không bất biến với sự thay đổi ánh sáng. Có một số nghiên cứu để tìm ra các đặc trưng ít nhạy với ánh sáng như [101,102] dựa trên mô hình trường Markov ngẫu nhiên nhằm thực hiện nhận dạng một số mẫu chất liệu nhưng cũng chỉ trong trường hợp thay đổi ánh sáng rất hạn chế (trong môi trường thí nghiệm). Để phát hiện được các mẫu chất liệu có cấu trúc bề mặt mịn và có sự thay đổi ánh sáng trong điều kiện tự nhiên phù hợp với ngữ cảnh bài toán của luận án đặt ra, phần này chúng tôi đề xuất kỹ thuật phát hiện mẫu chất liệu dựa vào đặc trưng nhiễu chất liệu. Giải pháp cho bài toán phát hiện mẫu chất liệu gồm hai pha [31]: 1) xây dựng mô hình cho mẫu chất liệu, hay nói khác hơn tìm được những thuộc tính đặc tả cho mẫu chất liệu; 2) dò tìm trong ảnh cho trước có mẫu chất liệu được đặc trưng bởi mô tả đó trên cơ sở đối sánh từng đặc trưng. Các mẫu chất liệu khác nhau sẽ có mô tả nhiễu khác nhau và việc phát hiện mẫu chất liệu có thể thực hiện bằng cách tìm kiếm mẫu nhiễu dư tương quan với mô tả nhiễu chất liệu đã xây dựng được. Kết quả thực nghiệm cho thấy kỹ thuật đề xuất phát hiện được các mẫu chất liệu trong trường hợp ánh sáng thay đổi lớn khá hiệu quả.

### **3.1. Biểu diễn mẫu chất liệu dựa đặc trưng nhiễu**

Vì nhiễu chất liệu phụ thuộc chủ yếu vào bản thân chất liệu hay nói khác hơn mỗi loại chất liệu khác nhau sẽ tạo ra đặc trưng nhiễu khác nhau. Song thể hiện ảnh

của chất liệu cũng phụ thuộc vào điều kiện thu nhận đặc biệt là ánh sáng. Do đó, có thể biểu diễn mẫu nhiễu dựa vào tập nhiễu của các ảnh chất liệu. Mẫu nhiễu chất liệu được xây dựng trên tập các ảnh nhiễu chất liệu theo quy tắc học tích lũy dần tạo ra đặc trưng nhiễu chất liệu, đặc trưng này ít nhạy với sự thay đổi ánh sáng.

### 3.1.1. Đặc trưng nhiễu chất liệu

Trong trường hợp mẫu chất liệu không có mô hình biểu diễn, mẫu chất liệu có thể là một mô hình chất liệu bao gồm một tập các ảnh chất liệu  $\mathcal{M} = \{\mathcal{M}_k\}_{k=1}^{N_p}$ ,  $N_p$  là số ảnh chất liệu dùng để xây dựng mẫu chất liệu. Đặc trưng nhiễu chất liệu  $n(\mathcal{M})$  có thể được tính toán bằng cách lấy trung bình của nhiễu phần tử ảnh (pixels)  $p^{(k)}$ ,  $k=1, \dots, N_p$ . Để tốc độ tính toán nhanh hơn, cần phải thực hiện các phép khử nhiễu từ bức ảnh trước khi lấy trung bình nhờ sử dụng bộ lọc nhiễu  $\mathcal{F}$  và lấy trung bình lượng nhiễu  $n^{(k)}$  như sau:

$$n^{(k)} = p^{(k)} - \mathcal{F}(p^{(k)}) \quad (3.1)$$

Có nhiều bộ lọc nhiễu [18,68] được đánh giá tốt cho các ảnh là cảnh tự nhiên, nhưng luận án đã sử dụng bộ lọc nhiễu wavelet [73] vì bộ lọc này cho kết quả lọc tốt nhất trong trường hợp các mẫu chất liệu tự nhiên.

### 3.1.2. Xây dựng đặc trưng nhiễu cho mô hình mẫu chất liệu

Với mỗi mẫu chất liệu dưới dạng mô hình  $\mathcal{M}$  gồm một tập  $N_p$  bức ảnh  $\{\mathcal{M}_k\}_{k=1}^{N_p}$  ( $N_p > 50$ ), ta thực hiện lần lượt các bước sau:

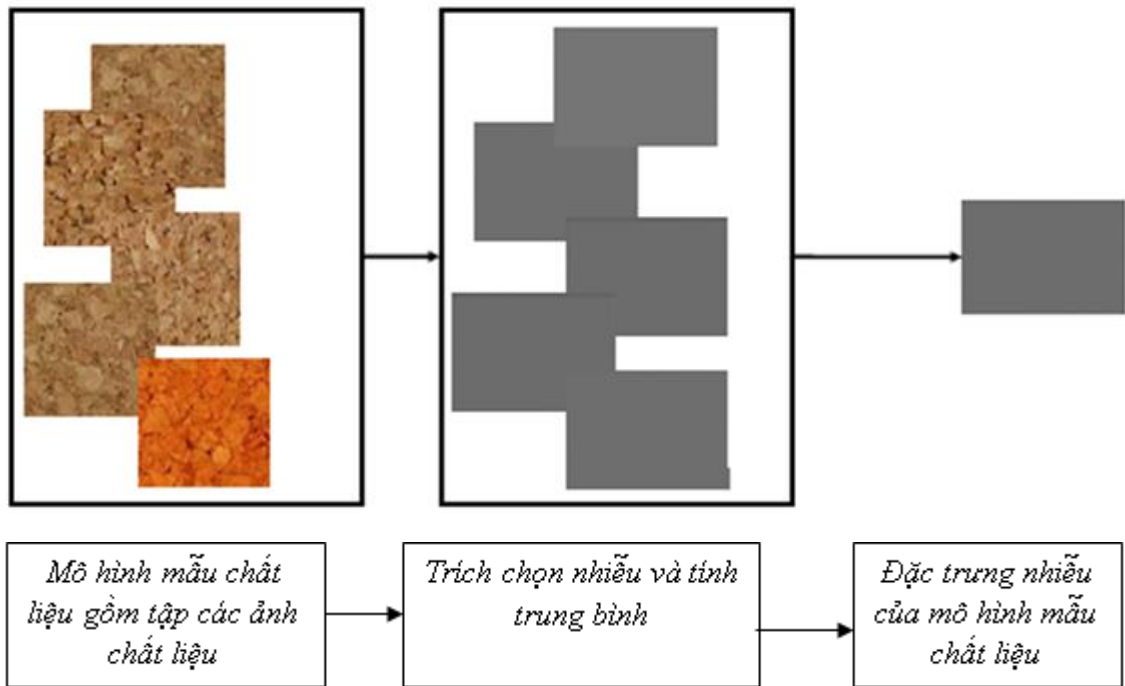
- Tính nhiễu cho tất cả các bức ảnh của cùng một mẫu chất liệu bằng hàm khử nhiễu  $\mathcal{F}$  (sử dụng bộ lọc khử nhiễu wavelet) để được các ảnh sau khử nhiễu  $\mathcal{F}(\mathcal{M}_k)$ .
- Tính nhiễu của từng ảnh  $n(\mathcal{M}_k)$  như sau:

$$n(\mathcal{M}_k) = \mathcal{M}_k - \mathcal{F}(\mathcal{M}_k) \quad (3.2)$$

- Tính trung bình tất cả các lượng nhiễu, thu được nhiễu của mẫu chất liệu  $\mathcal{M}$  như sau:

$$P_{\mathcal{M}} = \frac{\sum_{k=1}^{N_p} n(\mathcal{M}_k)}{N_p} \quad (3.3)$$

Do quá trình thu thập dữ liệu của mẫu chất liệu phải được thực hiện trước khi tạo mẫu nhiễu chất liệu với lượng dữ liệu lớn tại một thời điểm thường tốn thời gian rất lớn, để quá trình tạo đặc trưng nhiễu của mẫu chất liệu được thực hiện cùng với quá trình thu thập dữ liệu chất liệu theo thời gian. Kỹ thuật đề xuất dựa trên giải pháp tạo đặc trưng nhiễu theo phương pháp tích lũy dần như hình 3.1. Ban đầu, chỉ tạo nhiễu với một số ít mẫu dữ liệu các bức ảnh chất liệu. Sau đó với mỗi ảnh chất liệu thu nhận được, ta tiếp tục thực hiện quá trình tạo đặc trưng nhiễu chất liệu với nhiễu của chất liệu đưa vào và đặc trưng nhiễu đã có sẵn.



Hình 3.1. Mô hình xây dựng đặc trưng nhiễu cho mẫu chất liệu

Với các nhiễu chất liệu mới thêm vào, quá trình tính toán được thực hiện tương tự, cho đến khi số bức ảnh  $N_p$  đủ lớn ( $N_p > 50$ ). Giả sử có  $N_p$  ảnh chất liệu ban đầu.



Ta có đặc trưng nhiễu tính được:  $P_{\mathcal{M}} = \frac{\sum_{k=1}^{N_p} n(\mathcal{M}_k)}{N_p}$

Khi đó, nếu có một ảnh chất liệu  $T$  bất kỳ. Gọi  $P_{\mathcal{M}T}$  là đặc trưng nhiễu sẽ được tạo sau khi thêm  $T$ . Ta có công thức được biến đổi:

$$P_{\mathcal{M}T} = \frac{1}{N_p + 1} \left( \sum_{i=1}^{N_p} n(\mathcal{M}_i) + n(T) \right)$$

$$\Leftrightarrow (N_p + 1)P_{\mathcal{M}T} = \sum_{i=1}^{N_p} n(\mathcal{M}_i) + n(T) \Leftrightarrow (N_p + 1)P_{\mathcal{M}T} = N_p P_{\mathcal{M}} + n(T)$$

$$\Leftrightarrow P_{\mathcal{M}T} = \frac{1}{(N_p + 1)} \left( N_p P_{\mathcal{M}} + n(T) \right) \quad (3.4)$$

Đây chính là phép lấy trung bình các ảnh nhiễu chất liệu theo phương pháp tích lũy được sử dụng trong quá trình xây dựng đặc trưng nhiễu của mẫu chất liệu.

### 3.1.3. Thuật toán biểu diễn mẫu chất liệu dựa vào đặc trưng nhiễu RMBN

- **Vào:** Mẫu chất liệu  $\mathcal{M}$  (dưới dạng mô hình, gồm một tập ảnh chất liệu cùng kích thước  $m \times m$ ),  $\mathcal{M} = \{\mathcal{M}_i\}_{i=1}^{N_p}$ ,  $N_p$  là số ảnh chất liệu ( $N_p > 50$ ).
- **Ra:** Mô hình mẫu nhiễu chất liệu  $model\_noisy(\mathcal{M})$ , kích thước  $m \times m$
- Các bước thực hiện của thuật toán:

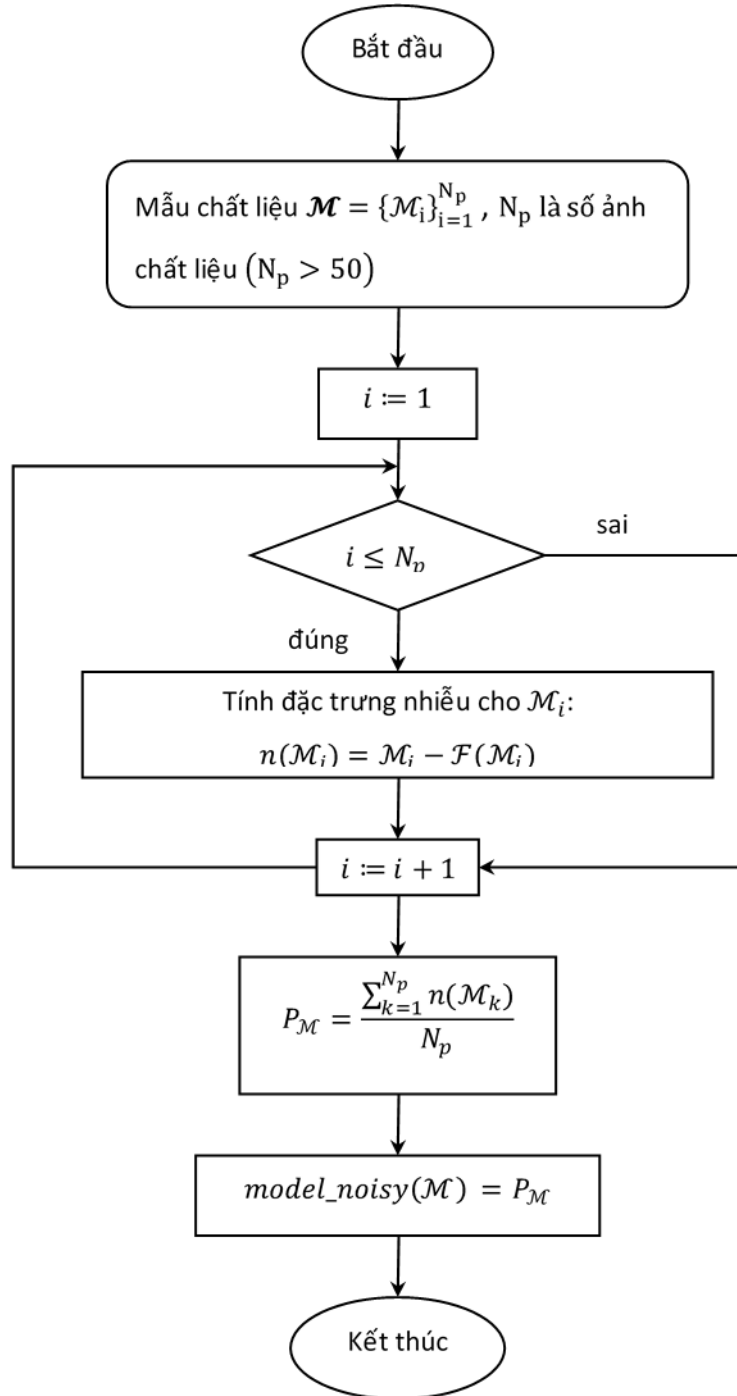
**Bước 1:** Với mỗi ảnh chất liệu  $\mathcal{M}_i$ ,  $i = 1, \dots, N_p$ . Tính đặc trưng nhiễu  $n(\mathcal{M}_i)$  theo công thức (3.2).

**Bước 2:** Tính  $model\_noisy(\mathcal{M}) = P_{\mathcal{M}}$  dựa vào công thức (3.3).

- Nhận xét:
  - Thuật toán RMBN luôn đảm bảo tính dừng và tính đúng đắn: Bước 1 và bước 2 của thuật toán chỉ đơn thuần là việc tính toán dựa vào các

công thức xác định (3.2) và (3.3). Do vậy thuật toán luôn dừng và luôn cho kết quả đúng đắn.

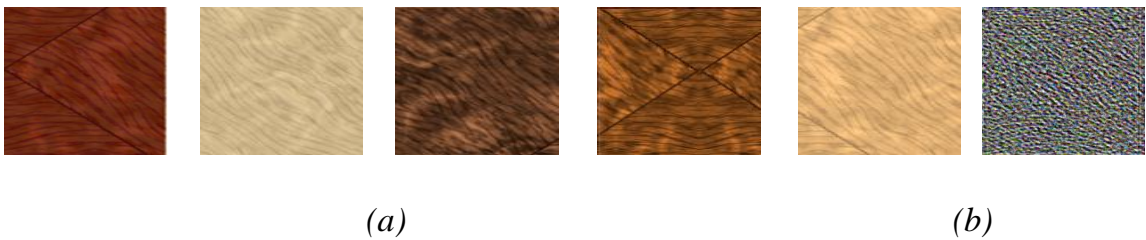
- Sơ đồ khối của thuật toán RMBN



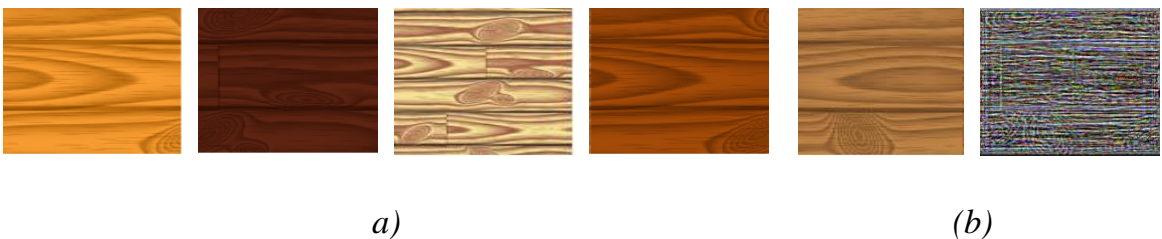
- Độ phức tạp tính toán của thuật toán: Tính độ phức tạp thuật toán dựa vào số mẫu chất liệu  $N_p$ , các tham số về mẫu chất liệu là cố định. Bởi vậy, độ phức tạp tính toán trên một ảnh chất liệu là  $O(1)$ . Vậy độ phức tạp thuật toán trong trường hợp xấu nhất là  $O(N_p)$ .
- Một số kết quả minh họa của thuật toán



Hình 3.2. Kết quả tìm đặc trưng nhiễu của mẫu gỗ 1 của thuật toán RMBN  
 (a) Mẫu chất liệu gỗ 1 dưới dạng mô hình gồm tập các chất liệu gỗ;  
 (b) Đặc trưng nhiễu của mẫu chất liệu gỗ 1



Hình 3.3. Kết quả tìm đặc trưng nhiễu của mẫu gỗ 2 của thuật toán RMBN  
 (a) Mẫu chất liệu gỗ 2 dưới dạng mô hình gồm tập các chất liệu gỗ;  
 (b) Đặc trưng nhiễu của mẫu gỗ 2



Hình 3.4. Kết quả tìm đặc trưng nhiễu của mẫu gỗ 3 của thuật toán RMBN  
 (a) Mẫu chất liệu gỗ 3 dưới dạng mô hình gồm tập các chất liệu gỗ;  
 (b) Đặc trưng nhiễu của mẫu chất liệu gỗ 3

### 3.2. Phát hiện mẫu chất liệu dựa vào đặc trưng mô hình nhiễu chất liệu

Dựa vào đặc trưng nhiễu của mẫu chất liệu dưới dạng mô hình được đề xuất ở trên, mục này luận án đề xuất thuật toán phát hiện mẫu chất liệu trong ảnh áp dụng tiếp cận phân lớp bằng phân phối Gauss.

#### 3.2.1. Phân lớp mẫu chất liệu dựa vào hàm phân phối Gauss

Thông thường xét theo tiếp cận thống kê thì các đối tượng thuộc cùng một lớp sẽ có cùng phân phối chuẩn, hay nói khác hơn độ tương quan của các cặp đối tượng cũng phải nằm trong một khoảng thống kê nhất định tuân theo một hàm phân phối cụ thể. Áp dụng tính chất này, luận án đã sử dụng hàm phân phối Gauss với tham số là độ tương quan giữa hai mẫu nhiễu để lựa chọn ngưỡng phân loại phù hợp với từng mẫu chất liệu khác nhau. Giả sử cần xác định ngưỡng phân lớp cho mẫu chất liệu  $\mathcal{M}$ . Ta xét cơ sở dữ liệu gồm tập đặc trưng nhiễu của các trường hợp có thể của mẫu chất liệu và tính lần lượt độ tương quan cho từng cặp. Giả sử có 2 đặc trưng nhiễu  $\mathcal{N}_1$  và  $\mathcal{N}_2$ , độ tương quan giữa  $\mathcal{N}_1$  và  $\mathcal{N}_2$  được tính như sau:

$$\rho(\mathcal{N}_1, \mathcal{N}_2) = \frac{(\mathcal{N}_1 - E[\mathcal{N}_1])(\mathcal{N}_2 - E[\mathcal{N}_2])}{\|\mathcal{N}_1 - E[\mathcal{N}_1]\| \|\mathcal{N}_2 - E[\mathcal{N}_2]\|} \quad (3.5)$$

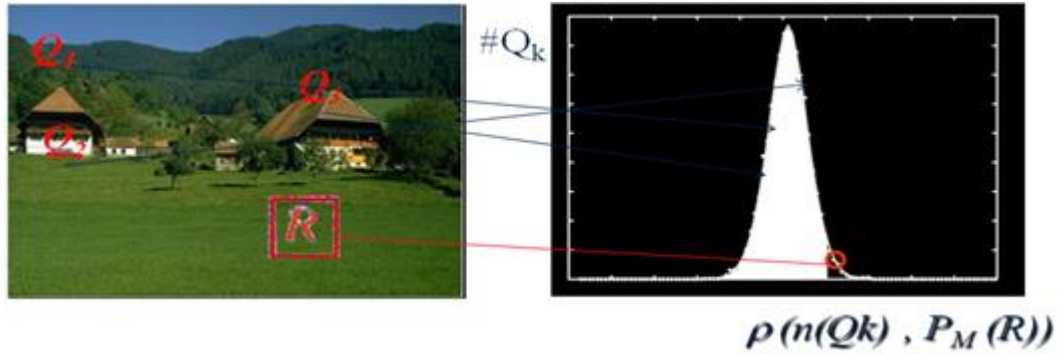
Trong đó,  $E[.]$  là phép lấy kỳ vọng. Nếu có được cơ sở dữ liệu các giá trị  $\rho_i$ , áp dụng vào hàm phân phối Gauss để xác định được ngưỡng phân loại  $\rho$  cụ thể cho từng mẫu chất liệu.

Tiếp theo là việc phân loại mẫu chất liệu trên ảnh như thế nào? Giả sử cần xét là một khối  $\mathfrak{R}$  trong ảnh  $\mathcal{J}$  (kí hiệu  $\mathcal{J}_{\mathfrak{R}}$ ) phải kiểm tra xem  $\mathfrak{R}$  có tương thích với đặc trưng nhiễu của mẫu chất liệu  $\mathcal{M}$  hay không, ta tính toán độ tương quan giữa số nhiễu  $n(\mathfrak{R}) = \mathcal{J}_{\mathfrak{R}} - \mathcal{F}(\mathfrak{R})$  với mẫu nhiễu chất liệu giả sử là  $P_{\mathcal{M}}$  như sau :

$$\rho(n(\mathfrak{R}), P_{\mathcal{M}}) = \frac{(n(\mathfrak{R}) - E[n(\mathfrak{R})])(P_{\mathcal{M}} - E[P_{\mathcal{M}}])}{\|n(\mathfrak{R}) - E[n(\mathfrak{R})]\| \|P_{\mathcal{M}} - E[P_{\mathcal{M}}]\|}$$

Nhiều sinh ra do thiết bị thu nhận cùng với nhiễu chất liệu tạo nên một đặc trưng nhiễu bất biến cho các loại mẫu chất liệu khác nhau. Nếu đặc trưng nhiễu của

hai loại mẫu chất liệu có độ tương quan lớn thì cũng có nghĩa rằng hai chất liệu đó tương tự nhau. Nói cách khác, nhiễu của chất liệu cần dò tìm phải nằm trong lân cận nhiễu của mẫu chất liệu cho trước, nếu chất liệu đó tương tự với mẫu chất liệu. Hình 3.5 minh họa trực quan cách chọn ngưỡng phân loại dựa vào hàm phân phối Gauss.



Hình 3.5. Minh họa vùng chất liệu  $R$  được chọn dựa vào phân phối Gauss

Ta có thể đánh giá độ tương quan  $\rho(n(Q), P_M(R))$  của khối  $Q$  và  $R$  có giống nhau hay không. Để giải quyết khó khăn này luận án sử dụng cách đánh giá bằng phân phối Gauss. Điểm thuận lợi ở đây là có thể thu được số lượng các mẫu chất liệu lớn, cần thiết để đánh giá các tham số  $\rho(n(Q), P_M(R))$ .

Đầu tiên, tính toán độ tương quan  $\rho(n(Q_k), P_M(R))$  giữa các vùng trên ảnh vào  $Q_k, k = 1 \dots N_R$  với mô hình đặc trưng nhiễu đã xác định với kích thước giống nhau. Ta tính toán các  $\rho(n(Q_k), P_M(R)), k = 1 \dots N_R$ . Quá trình ra quyết định được xác định thông qua phân phối Gauss (Generalized Gaussian distribution) với hàm tích lũy  $G(x)$ . Thông qua mô hình phân phối Gauss, sẽ xác định được  $\rho$  hợp lý để lựa chọn kết quả phát hiện chính xác. Áp dụng mô hình này có thể tổng quát hoá một phân phối Gauss biến đổi ngẫu nhiên với ước lượng phân loại sẽ thu được giá trị:

$$\varepsilon = G(\rho(n(Q), P_M(R))) \quad (3.6)$$

### 3.2.2. Thuật toán phát hiện mẫu chất liệu dựa vào nhiễu DMBNF

- Vào: Mô hình của mẫu nhiễu chất liệu  $\mathcal{M}$  :  $model\_noisy(\mathcal{M})$  có kích thước  $m \times m$  và ảnh cần tìm mẫu chất liệu  $\mathcal{J}$ .

- **Ra:** Tập  $R$  gồm các vùng chứa mẫu chất liệu  $\mathcal{M}$  trong ảnh  $\mathcal{J}$ .
- Các bước thực hiện của thuật toán:

**Bước 1:** Tính ảnh nhiễu  $\mathcal{J}'$  của ảnh  $\mathcal{J}$  dựa vào công thức (3.2). Chọn ngưỡng phân loại  $\varepsilon$ . Khởi tạo  $R := \{\emptyset\}$ .

**Bước 2:** Phân hoạch ảnh  $\mathcal{J}'$  thành tập các ảnh con  $RI' = \{\mathfrak{R}_i\}_{i=1}^n$  có cùng kích thước  $m \times m$ .

**Bước 3:** Với mỗi  $\mathfrak{R}_i \in RI', i = 1, \dots, n$ , thực hiện:

**Bước 3.1:** Tính độ tương quan giữa  $\mathfrak{R}_i$  và  $model\_noisy(\mathcal{M})$  dựa vào (3.5), ta được  $\rho_{\mathfrak{R}_i}$ .

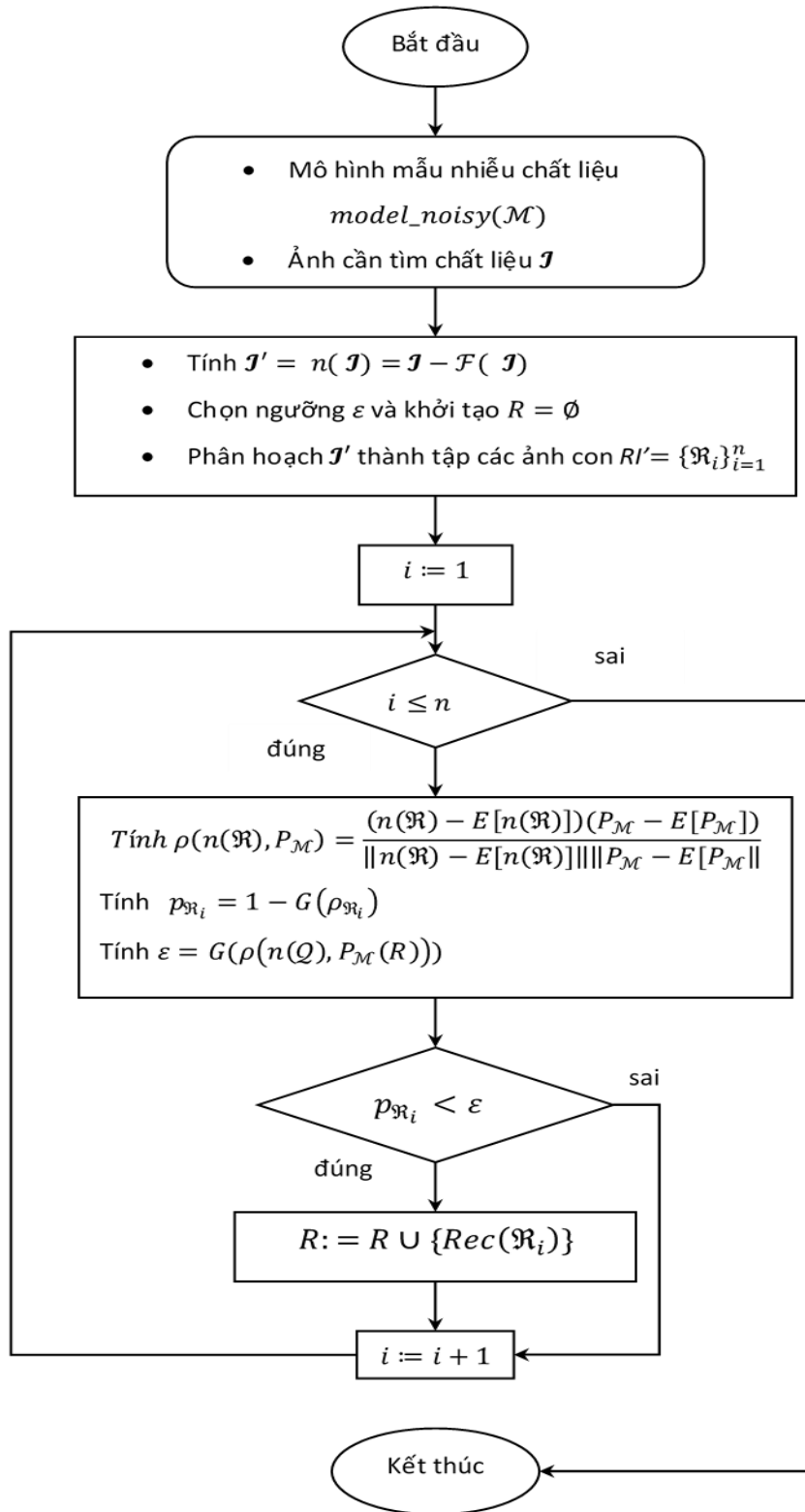
**Bước 3.2:** Áp dụng phân phối Gauss cho giá trị  $\rho_{\mathfrak{R}_i}$  ta được  $G(\rho_{\mathfrak{R}_i})$ , sau đó tính giá trị  $p_{\mathfrak{R}_i} = 1 - G(\rho_{\mathfrak{R}_i})$

**Bước 3.3:** Kiểm tra nếu  $p_{\mathfrak{R}_i} < \varepsilon$  ( $\varepsilon$  được xác định dựa vào công thức (3.6)) thì kết luận  $\mathfrak{R}_i$  chính là vùng ảnh con chứa mẫu chất liệu  $\mathcal{M}$  và lưu lại  $Rec(\mathfrak{R}_i)$ , tức là  $R := R \cup \{Rec(\mathfrak{R}_i)\}$ . Ngược lại,  $\mathfrak{R}_i$  không phải vùng chứa mẫu chất liệu cần tìm.

- Nhận xét:
  - Thuật toán DMBNF luôn đảm bảo tính dừng: Tập  $RI'$  là hữu hạn, do đó thuật toán sẽ dừng trong bước 3.
  - Thuật toán DMBNF luôn đảm bảo tính đúng đắn: Vì  $\varepsilon$  được xác định trước, do đó  $p_{\mathfrak{R}_i} < \varepsilon$  luôn xác định được chính xác trong bước 3, nghĩa là kết quả trả về của thuật toán là đúng đắn.
- Độ phức tạp tính toán của thuật toán DMBNF

Tính toán độ phức tạp thuật toán dựa trên kích thước ảnh đầu vào là  $n \times n$ . Việc tính nhiễu có độ phức tạp là  $O(n^2)$ , thực hiện duyệt tất cả các cửa sổ có độ phức tạp  $O(n^2)$ . Kích thước mẫu chất liệu là cố định nên việc tính toán so sánh tại một cửa sổ có độ phức tạp là hằng số. Do vậy, độ phức tạp thuật toán là  $O(n^2)$ .

- Thuật toán DMBNF được biểu diễn dưới dạng sơ đồ khối:



### 3.2.3. Thử nghiệm

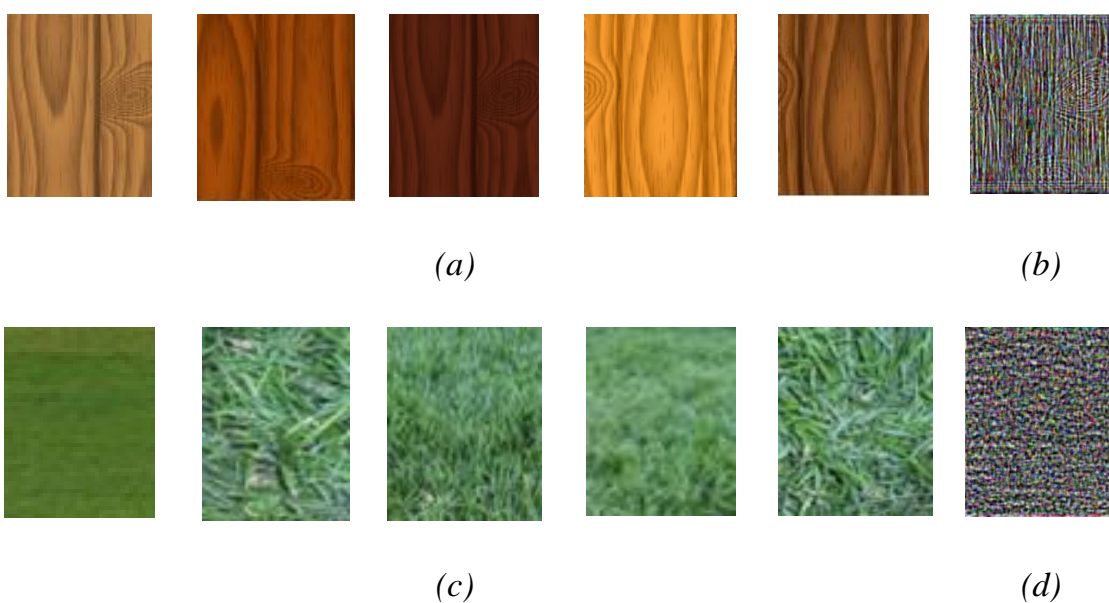
Luận án đã cài đặt thử nghiệm cho thuật toán DMBNF được đề xuất ở trên, sử dụng ngôn ngữ Visual C++ 9.0 (*Chi tiết xem phụ lục A3 và A4*), với cấu hình máy Intel Pentium dual core = 2.4, 2GB RAM. Việc xây dựng mẫu nhiễu cho từng loại chất liệu, mỗi loại chất liệu chúng tôi tính toán trung bình nhiều trên 50 mẫu chất liệu kích thước  $m$  được chọn theo kinh nghiệm (trong thực nghiệm này  $m$  chọn từ 180 đến 250) phụ thuộc vào từng loại chất liệu riêng, với các điều kiện môi trường khác nhau. Chúng tôi thử nghiệm với cơ sở dữ liệu cũng được trích rút từ tập dữ liệu MIT-CSAIL và dữ liệu thu thập được như đã đề cập trong phần thử nghiệm của thuật toán phát hiện mẫu chất liệu dựa vào đặc trưng bất biến địa phương sử dụng các ảnh có sự thay đổi ánh sáng với các mẫu chất liệu gỗ, cỏ, da và tóc. Với mỗi loại chất liệu chúng tôi lấy 50 ảnh kích thước thay đổi từ 180 đến 250 tùy thuộc mỗi loại chất liệu khác nhau để xây dựng mẫu nhiễu cho chất liệu đó. Chúng tôi cũng thử nghiệm trên 1470 ảnh vào có chất liệu gỗ, 2014 ảnh vào có chất liệu cỏ, 693 ảnh vào có chất liệu da và 818 ảnh vào có chất liệu tóc. Thực hiện thuật toán với thông số ngưỡng  $\varepsilon = 2$  được đánh giá trong bảng 3.1. Kết quả thực nghiệm cho thấy thuật toán DMBNF phát hiện khá hiệu quả cho các trường hợp mẫu chất liệu có sự thay đổi ánh sáng lớn và cho các mẫu chất liệu có tính kết cấu bề mặt mịn như da. Trong khi đó mẫu chất liệu gỗ hay cỏ thì độ chính xác không tăng lên nhiều so với kỹ thuật phát hiện mẫu chất liệu dựa vào đặc trưng bất biến địa phương.

**Bảng 3.1. Kết quả đánh giá thuật toán DMBNF**

Mẫu chất liệu thử nghiệm	Kết quả phát hiện mẫu chất liệu bằng thuật toán DMBNF				
	Số ảnh	Đúng	Sai	Tỉ lệ sai	Độ chính xác
Gỗ	1470	1455	15	1,00	98,97%
Cỏ	2014	2001	13	0,60	99,35%
Da	693	647	46	6,60	93,33%
Tóc	818	725	93	11,36	88,63%



Dưới đây là một số hình ảnh minh họa. Hình 3.6 minh họa mẫu nhiễu của chất liệu gỗ và chất liệu cỏ được tạo ra từ một tập các ảnh chất liệu gỗ và cỏ. Qua đó, ta thấy cách tiếp cận này không những phân biệt các loại chất liệu khác nhau, mà còn có thể phân biệt được các kiểu của một mẫu chất liệu. Hình 3.7 và 3.8 minh họa một số kết quả của kỹ thuật đề xuất, hình 3.7 minh họa cho việc phát hiện mẫu chất liệu cỏ trong ảnh vào có mẫu chất liệu cỏ và bị thay đổi ánh sáng lớn (mẫu chất liệu được lấy ra trong clip thứ nhất và ảnh vào được trích ra trong clip thứ ba của tập dữ liệu MIT-CSAIL). Hình 3.8 minh họa kết quả phát hiện mẫu chất liệu gỗ trong ảnh vào cũng có sự thay đổi ánh sáng, và kết quả phát hiện được khá chính xác.



*Hình 3.6. Các mẫu nhiễu của một số chất liệu  
 (a) Mẫu chất liệu gỗ; (b) Đặc trưng nhiễu của mẫu gỗ;  
 (c) Mẫu chất liệu cỏ; (d) Đặc trưng nhiễu của mẫu cỏ*



(a)



(b)



(c)



(d)



(e)



(f)

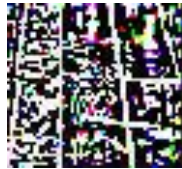
Hình 3.7. Kết quả phát hiện mẫu chất liệu cỏ cây bằng thuật toán DMBNF  
(a) Mẫu chất liệu cỏ cây; (b) Ảnh vào; (c) Đặc trưng nhiễu của chất liệu cỏ cây;  
(d) Ảnh nhiễu của ảnh vào (e) Kết quả phát hiện trên ảnh nhiễu vào;  
(f) Kết quả phát hiện trên ảnh vào



(a)



(b)



(c)



(d)



(e)



(f)

Hình 3.8. Kết quả phát hiện mẫu chất liệu gỗ bằng thuật toán DMBNF  
(a) Mẫu chất liệu gỗ; (b) Ảnh vào; (c) Đặc trưng nhiễu của mẫu chất liệu gỗ;  
(d) Nhiễu của ảnh vào; (e) Kết quả phát hiện trên ảnh nhiễu;  
(f) Kết quả phát hiện trên ảnh vào

### 3.3. Kết luận chương 3

Ảnh thu được sau quá trình thu nhận ảnh thường không tránh khỏi nhiễu gây ra bởi bản thân đối tượng tạo nên và nhiễu do môi trường. Mỗi chất liệu khác nhau

sẽ có một lượng nhiễu khác nhau đặc trưng cho chất liệu đó, lượng nhiễu đặc trưng cho mỗi chất liệu được gọi là nhiễu chất liệu. Dựa vào đặc điểm này luận án đã đề xuất một mô hình nhiễu chất liệu, trên cơ sở đó đưa ra thuật toán phát hiện mẫu chất liệu dựa vào đặc trưng nhiễu sử dụng tính toán độ tương quan giữa các vùng ảnh và mẫu nhiễu của chất liệu cho trước. Kỹ thuật đề xuất đã được cài đặt và thử nghiệm trên một số mẫu chất liệu như gỗ, cỏ, da và tóc. Kết quả đạt được cho thấy kỹ thuật đề xuất khá hiệu quả đối với các mẫu chất liệu có sự tương đồng về cấu trúc hình học của bề mặt và cả khi ánh sáng bị thay đổi.

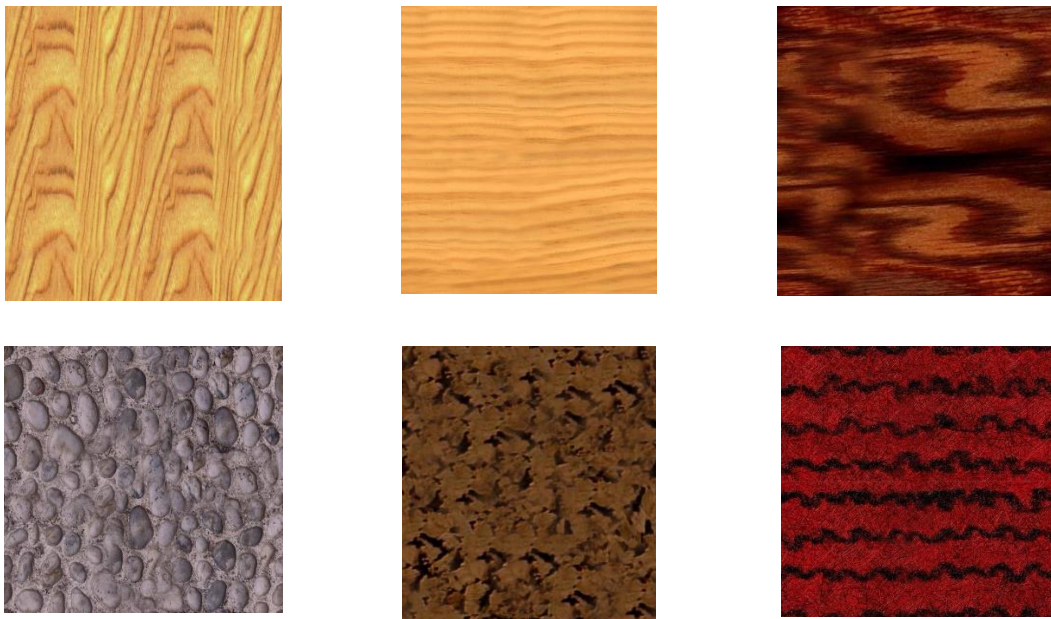
## **Chương 4. PHÁT HIỆN MẪU CHẤT LIỆU DỰA VÀO HÌNH HỌC FRACTAL**

Mẫu chất liệu trong tự nhiên thường có cấu trúc lặp lại tại các vùng khác nhau, và có mối quan hệ không gian phân biệt như gỗ, thảm, vải v.v. Cấu trúc lặp lại cùng với các mối quan hệ không gian nhất định sẽ tạo nên các cấu trúc quang học lẫn cấu trúc hình học sẽ đặc tả được nhiều mẫu chất liệu khác nhau trong cuộc sống. Có một số nghiên cứu tìm cấu trúc lặp lại về mặt hình học của hình dáng đối tượng trong ngữ cảnh nhận dạng và tra cứu ảnh như [43,114]. Fractal là công cụ toán học mạnh mẽ để đặc tả các vùng lặp lại chi tiết cho mỗi một cấu trúc hình học và quang học của mẫu chất liệu [34]. Hơn nữa, biểu diễn và phát hiện mẫu chất liệu dựa vào hình học Fractal sẽ giải quyết được trong trường hợp mẫu chất liệu xuất hiện trong ảnh có sự thay đổi afin, đặc biệt là vấn đề tỉ lệ toàn cục [5]. Trong chương này luận án đề xuất một phương pháp phát hiện mẫu chất liệu có cấu trúc lặp lại các chi tiết trên cơ sở biểu diễn Fractal dựa trên hệ hàm lặp IFS (Iterated Function System) của hình học Fractal.

### **4.1. Đặt vấn đề**

Một mô tả mẫu chất liệu tốt cho việc biểu diễn và phát hiện chất liệu là phải có khả năng thể hiện được các thuộc tính trực quan của cấu trúc chất liệu nhưng cũng phải bất biến đối với các biến đổi của môi trường như thay đổi độ chiếu sáng, hay các phép biến đổi hình học như thay đổi tỉ lệ, quay, dịch chuyển và các dạng méo mó khác. Việc tìm kiếm các đặc trưng bất biến đã bắt đầu từ thế kỷ 19 [13]. Khi đó, các nhà nghiên cứu sử dụng các kiến thức toán học về đại số và bất biến hình học, đã xây dựng được các nghiên cứu định lượng về các bất biến đối với điểm nhìn từ một bức ảnh. Khi đó các tác giả cũng đưa ra một số mô tả bất biến điểm nhìn và đã ứng dụng được vào bài toán nhận dạng. Có nhiều phương pháp trích chọn và xây dựng đặc trưng bất biến có thể mô tả được chất liệu như SIFT, LBP, CS-LBP v.v. Tuy nhiên, chúng vẫn chỉ áp dụng được trên một số chất liệu không đồng nhất về mức xám, cấu trúc bề mặt không đồng đều. Phần này luận án cũng đề xuất một kỹ thuật khác để

phát hiện chất liệu, đặc biệt là chất liệu có cấu trúc. Như chúng ta đã biết, Fractal là một mô hình toán học cho phép mô tả các hiện tượng có sự lặp lại của trong tự nhiên [115]. Các chất liệu có cấu trúc hầu hết đều là các bề mặt chất liệu tự nhiên, luôn có cấu trúc lặp lại tại một hoặc một số phần của ảnh. Hình 4.1 minh hoạ các chất liệu có cấu trúc và tính chất lặp lại của chúng. Trên cơ sở đó, luận án tận dụng tính chất lặp lại của các chất liệu có cấu trúc để tìm ra thành phần chính đặc trưng cho cả mẫu chất liệu đó, nghĩa là sẽ thu được một số phần tử đặc biệt, mỗi phần tử đó sẽ biểu diễn một thành phần cốt lõi cô đọng nhất của mẫu chất liệu đó chính là cách tiếp cận dựa trên lý thuyết hàm lặp của Fractal. Luận án đề xuất một phương pháp phát hiện mẫu chất liệu dựa trên nền tảng lý thuyết hàm lặp, xấp xỉ các ánh xạ co giữa các vùng giống nhau trong ảnh, hay nói khác hơn nó được biểu diễn bằng các thành phần cô đọng nhất của một chuỗi các hàm lặp IFS (Iterated Function System), đó chính là biểu diễn Fractal của mẫu chất liệu.



*Hình 4.1. Tính chất lặp lại của các mẫu chất liệu*

## 4.2. Cơ sở toán học

Có thể nói cùng với lý thuyết topo, hình học Fractal cung cấp một công cụ khảo sát các tính chất của các sự vật, hiện tượng trong tự nhiên vô cùng mạnh mẽ. Vật lý và toán học thế kỷ 20 đối đầu với sự xuất hiện của tính hỗn độn trong nhiều quá trình có tính quy luật của tự nhiên. Từ sự đối đầu đó hình thành một lý thuyết mới, chuyên nghiên cứu về các hệ phi tuyến, gọi là lý thuyết hỗn độn. Việc khảo sát bằng các bài toán phi tuyến đòi hỏi rất nhiều công sức tính toán và thể hiện các quan sát một cách trực quan. Tuy vậy, sự phát triển của lý thuyết này bị hạn chế. Gần đây với sự ra đời của lý thuyết fractal và sự hỗ trợ đắc lực của khoa học máy tính, các nghiên cứu chi tiết về sự hỗn độn được đẩy mạnh. Vai trò hình học fractal trong lĩnh vực này thể hiện một cách rõ ràng hành vi kỳ dị của các tiến trình được khảo sát, qua đó tìm ra được các đặc trưng hoặc các cấu trúc tương tự nhau trong các ngành khoa học khác nhau. Hình học fractal đã được áp dụng vào nghiên cứu lý thuyết từ tính, lý thuyết các phức chất trong hóa học, lý thuyết tái định chuẩn, các nghiệm của các hệ phương trình phi tuyến (được giải dựa trên phương pháp xấp xỉ liên tiếp của Newton trong giải tích số) [91]. Gần đây lý thuyết hình học Fractal đã dần len lỏi vào lĩnh vực khoa học máy tính như nén ảnh [96], nhận dạng mẫu [75,76], nhận biết kí tự trong xác thực mật khẩu [74,75] v.v.

Phần này luận án đề xuất vận dụng hình học Fractal để biểu diễn và phát hiện các mẫu chất liệu có tính lặp lại tại các chi tiết trong thực tế dựa trên nguyên lý nén ảnh Fractal đã được nghiên cứu rất mạnh mẽ trong thời gian qua. Theo cách tiếp cận này, mẫu chất liệu có thể được biểu diễn bằng tập các khối ảnh con có kích thước nhỏ hơn rất nhiều so với ảnh gốc cùng tập các phép biến đổi tương ứng nhằm từ các phép biến đổi này có thể tác động nhiều lần lên các khối ảnh con khác ( bị thay đổi bởi các phép biến đổi hình học afin và ánh sáng) để tạo sinh các mẫu chất liệu xấp xỉ với mẫu chất liệu ban đầu, nếu hệ số xấp xỉ tìm được nhỏ hơn một ngưỡng cho phép thì có thể kết luận khối ảnh con và khối ảnh con khác đó là giống nhau. Sở dĩ phải thực hiện điều này nhằm khắc phục được phần nào vấn đề khó khăn của việc so sánh hai thể

hiện khác nhau (do điều kiện thu nhận khác nhau) của cùng một mẫu chất liệu khi không biết trước điều kiện thu nhận trong xử lý ảnh và thị giác máy. Cơ sở toán học của biểu diễn và phát hiện mẫu chất liệu dựa vào hình học Fractal gồm Lý thuyết hàm lặp, ánh xạ co trên không gian metric đầy đủ, điểm bất động v.v đã được trình bày trong chương 1. Biểu diễn Fractal của mẫu chất liệu dựa trên phương pháp tìm hệ hàm lặp IFS (Iterated Function Systems). IFS như đã định nghĩa trong chương 1 đó là một tập các ánh xạ trong không gian 3D (vị trí và giá trị cường độ của điểm ảnh).

Gọi không gian metric trên tập ảnh số là một cặp  $(M, d)$ .  $M$  là tập ảnh,  $d$  là một độ đo. Để biểu diễn ảnh  $M_1 \in M$ , ta phải tìm ánh xạ  $w: M \rightarrow M$  sao cho :

$\exists 0 < c < 1, \forall \mu, \vartheta \in M; d(\mu, w(\vartheta)) \leq c \cdot d(\mu, \vartheta)$ ,  $\mu, \vartheta$  là các khối con của ảnh  $M_1$  thì điều kiện phân hoạch các khối con của  $M_1$  như sau :

$$\forall \mu_i \in M_1, i = \overline{1, n}, \cup \mu_i = M_1, \mu_i \cap \mu_j = \emptyset$$

Một mô tả IFS  $\mathcal{W}$  được định nghĩa bằng một tập con hữu hạn các đặc tả ảnh :

$$\mathcal{W} = \{w_0, w_1, \dots, w_{n-1}\}$$

Gọi  $H(M)$  là một thành phần của  $M$ , nếu  $\exists \mathfrak{R} \in H(M)$  thì :

$$\mathcal{W}\mathfrak{R} \in H(M), \mathcal{W}\mathfrak{R} = w_1\mathfrak{R}, \dots, w_{n-1}\mathfrak{R}$$

Khi đó toán tử co  $\mathcal{W}$  trong không gian metric đầy đủ mới  $H(M)$  sẽ tồn tại điểm cố định  $F(\mathcal{W}) = \lim_{n \rightarrow \infty} \mathcal{W}^n \mathfrak{R}, \mathfrak{R} \in H(M)$ .

Như vậy với một khối chất liệu  $\mathfrak{R}$  bất kì nào đó, có thể trích rút được một tập các đặc tả  $\mathcal{W}$  làm đặc trưng riêng cho mẫu chất liệu. Đặc trưng này sẽ bất biến với các phép biến đổi ảnh, đặc biệt là các phép biến đổi afin.

Vấn đề là chúng ta phải trích rút  $\mathcal{W}$  như thế nào. Có nhiều phương pháp trích rút  $\mathcal{W}$  thông qua các phương pháp xấp xỉ như mạng nơron nhân tạo, các phương pháp học máy v.v.  $\mathcal{W}$  chính là một trong những thành phần biểu diễn chất liệu có cấu trúc lặp lại chính xác tại các tỉ lệ khác nhau.



### 4.3. Biểu diễn mẫu chất liệu dựa vào hình học Fractal

#### 4.3.1. Đặc trưng hình học Fractal cho chất liệu

Đặc trưng Fractal của chất liệu chính là các khối ảnh chất liệu con cùng với phép biến đổi affine xấp xỉ tương ứng và từ đó có thể xây dựng lại được mẫu chất liệu ban đầu thông qua các đặc trưng đó. Đặc trưng chất liệu trong cách tiếp cận của luận án dựa trên ý tưởng Fractal. Nó thể hiện tính chất lặp lại các chi tiết của mẫu chất liệu bằng các tham số của các phép biến đổi địa phương. Phần này luận án trình bày ngắn gọn cách thức xây dựng đặc trưng Fractal của mẫu chất liệu. Xét mẫu chất liệu  $X$  là một mảng 2 chiều  $m \times m$ . Ta xét một phân hoạch của  $X$  thành các khối con không giao nhau  $\mathfrak{R}_i$ ,  $1 \leq i \leq n$  sao cho  $X = \cup \mathfrak{R}_i$ , các  $\mathfrak{R}_i$  gọi là các khối dãy. Kết hợp với các khối dãy  $\mathfrak{R}_i$  là các khối miền lớn hơn  $\mathcal{D}_i \in X$  sao cho  $\mathfrak{R}_i \approx w_i \mathcal{D}_i$ , trong đó  $w_i$  là ánh xạ co 1-1. Giả sử hàm ảnh  $\mathcal{U}(\mathfrak{R}_i) = (\mathcal{U}|\mathfrak{R}_i)$  là một xấp xỉ của bản sao  $\mathcal{U}(\mathcal{D}_i) = (\mathcal{U}|\mathcal{D}_i)u(\mathcal{D}_i) = u|\mathcal{D}_i$  sao cho  $\mathcal{U}(\mathfrak{R}_i) = \Phi_i(\mathcal{U}(w_i(\mathcal{D}_i)))$ , trong  $\Phi_i : \mathfrak{R} \rightarrow \mathfrak{R}$  là ánh xạ affine mức xám, tức là  $\Phi_i = \alpha_t(t) + \beta_i$ , như vậy  $\mathcal{U}(x) \approx \mathcal{W}\mathcal{U}(x) = \alpha_i(\mathcal{U}(w_i(x))) + \beta_i, \forall x \in \mathfrak{R}_i$ .

$\mathcal{W}$  là biểu diễn Fractal của mẫu chất liệu, nếu toán tử  $\mathcal{W}$  là ánh xạ co của một hàm ảnh thích hợp trong không gian metric đầy đủ  $\mathcal{F}(X)$ , khi đó sẽ tồn tại một điểm bất động duy nhất  $\bar{u} \in \mathcal{F}(X)$  sao cho  $\mathcal{W}\bar{u} = \bar{u}$ . Với mỗi khối  $\mathfrak{R}_i$  chọn một  $\mathcal{D}_{i(j)}$  là xấp xỉ tốt nhất, nghĩa là sai số xấp xỉ bé nhất  $\Delta_{i,j} = \min_{\alpha, \beta \in \pi} \|\mathcal{U}(\mathfrak{R}_i) - \alpha\mathcal{U}(\mathcal{D}_{i(j)}) - \beta\|$ ,  $\alpha, \beta$  là không gian tham số. Thực tế, cực tiểu này tương đương với cực tiểu trong sai số của định lý Collage  $d(\mathcal{U}, \mathcal{W}\mathcal{U}) = \|\mathcal{U} - \mathcal{W}\mathcal{U}\|$ . Như vậy điểm bất động  $\bar{u}$  sẽ đạt được xấp xỉ  $\mathcal{U}$  đối với ảnh ban đầu.  $\bar{u}$  có thể được sinh ra từ một dãy các phép lặp bắt đầu từ  $\mathcal{U}_0, \dots, \mathcal{U}_{n+1} = \mathcal{W}\mathcal{U}_n$  và sẽ hội tụ đến  $\bar{u}$ . Phép biến đổi Fractal là một trường hợp của phép biến đổi toàn cục, vì một phần của ảnh được xấp xỉ bởi phần khác trong ảnh.

Vấn đề quan trọng ở đây là phải tìm  $\mathcal{W}$  sao cho lấy một ảnh đầu vào và nhận được một ảnh đầu ra gần nhất với ảnh đầu vào. Nhưng làm sao để chúng ta biết khi nào  $\mathcal{W}$  là phép biến đổi co, tức là chúng ta phải định nghĩa một khoảng cách hay một metric giữa hai ảnh.

- Metric trên ảnh

Một metric là một hàm đo khoảng cách giữa hai đối tượng nào đó. Ví dụ, các đối tượng có thể là hai điểm trên đường thẳng thực, metric có thể là giá trị tuyệt đối của sự sai khác giữa chúng. Ở đây luận án sử dụng khái niệm “metric”, không dùng “sự sai khác” hay “khoảng cách”, vì nó có ý nghĩa tổng quát, metric giữa hai điểm, hai tập hợp hay giữa hai ảnh. Có nhiều metric có thể được chọn để sử dụng, nhưng ở đây có thể dùng hai metric sau:

- Metric supremum:

$$d_{\text{sup}} = \sup_{(x,y) \in I^2} |f(x,y) - g(x,y)|$$

- Metric rms (root mean square):

$$d_{\text{rms}}(f, g) = \sqrt{\int_{I^2} f(x,y) - g(x,y) dx dy}$$

Đây là các metric đơn giản nhất, nhưng phù hợp đối với dữ liệu ảnh chất liệu.

#### 4.3.2. Xây dựng mô tả Fractal cho mẫu chất liệu

Nguyên tắc cơ bản để biểu diễn ảnh fractal là biểu diễn một bức ảnh bằng một hệ thống hàm lặp IFS, có điểm bất động của hàm này gần với ảnh ban đầu. Điểm bất động này được gọi là fractal của ảnh. Sau đó, mỗi IFS được mã hóa bằng một ánh xạ co với các hệ số. Định lý điểm bất động của Banach đã chứng tỏ rằng trong một không gian metric đầy đủ, điểm bất động của một ánh xạ có thể được khôi phục bởi cài đặt lặp với một phần tử ban đầu bất kỳ trong không gian đó. Vì vậy, quá trình mã hóa hay biểu diễn ảnh là phải tìm được một IFS có điểm bất động tương tự với ảnh đã cho. Cách tiếp cận dựa trên định lý Collage, định lý này đưa ra một ngưỡng khoảng

cách giữa ảnh được biểu diễn cô đọng và điểm cố định của IFS. Vì vậy, một phép biến đổi phù hợp được xây dựng bằng định lý Collage từ ảnh lên chính bản thân nó bằng một sai số đủ nhỏ (khoảng cách giữa Collage và ảnh) đảm bảo điểm cố định của phép biến đổi gần với ảnh gốc nhất.

Biểu diễn Fractal của mẫu chất liệu dựa trên nguyên lý IFS như sau:

- Phân hoạch mẫu chất liệu thành các khối dây  $\mathfrak{R}_i$  không giao nhau.
- Phân hoạch mẫu chất liệu thành các khối miền  $\mathcal{D}_i$  lớn hơn  $\mathfrak{R}_i$  (có thể chồng lên nhau).
- Với mỗi khối dây, tìm khối miền và phép biến đổi tương ứng sao cho nếu áp dụng phép biến đổi này lên khối miền thì kết quả thu được sẽ xấp xỉ tốt nhất với khối dây tương ứng.
- Lưu vị trí của khối dây, khối miền và các tham số của phép biến đổi tương ứng và đó chính là biểu diễn Fractal của chất liệu.

Trong đó, phân hoạch mẫu chất liệu thành tập các khối dây và các khối miền là một yếu tố quan trọng, nó phụ thuộc vào việc thử nghiệm trên từng loại ảnh theo từng kỹ thuật phân hoạch khác nhau điển hình trong xử lý ảnh có kỹ thuật phân hoạch cây tứ phân, đồng thời thử nghiệm trên các tham số về kích thước các phân hoạch cũng như mối tương quan giữa khối dây và khối miền. Mặt khác, để tìm được phép biến đổi  $\mathcal{W}_j$  sao cho khi tác động lên khối miền  $\mathcal{D}_j$  tức là  $\mathcal{W}_j(\mathcal{D}_j)$  thì  $\mathcal{W}_j(\mathcal{D}_j)$  phải là xấp xỉ tốt nhất với khối dây  $\mathfrak{R}_i$  tương ứng, và  $\mathcal{W}_j$  là phép biến đổi co như đã chứng minh trong chương 1, do vậy các khối miền  $\mathcal{D}_j$  phải lớn hơn các khối dây  $\mathfrak{R}_i$ .

Vấn đề chính của biểu diễn Fractal cho mẫu chất liệu là việc tìm ánh xạ  $\mathcal{W}$  tương ứng giữa khối miền và khối dây. Với mỗi khối dây, so sánh nó với mỗi khối miền được biến đổi, phép biến đổi này cũng được chứng minh phải là phép biến đổi affine. Phép biến đổi  $\mathcal{W}$  là tổ hợp của các phép biến đổi hình học và ánh sáng. Với

mỗi mẫu chất liệu mức xám  $\mathcal{M}$ , nếu kí hiệu  $z$  là cường độ pixel tại vị trí  $(x,y)$  thì  $\mathcal{W}$  có thể được biểu diễn:

$$\mathcal{W} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & s \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} e \\ f \\ o \end{bmatrix} \quad (4.1)$$

Các hệ số  $a, b, c, d, e, f$  là các hệ số hình học của phép biến đổi tỉ lệ, quay và dịch chuyển và  $s, o$  là các hệ số độ tương phản và độ sáng. Các phép biến đổi hình học trong luận án này chỉ hạn chế một số hướng quay, tỉ lệ và dịch chuyển nhất định. Nên việc so sánh khối miền và khối dãy tương ứng được thực hiện theo 3 bước: Áp dụng một trong số các hướng xác định lên khối miền  $\mathcal{D}_j$ , sau đó quay và thu nhỏ khối miền  $\mathcal{D}_j$  để được kích thước bằng khối dãy  $\mathcal{R}_i$  tương ứng rồi đối sánh với nhau. Bước cuối cùng ta tính toán các tham số  $s$  và  $o$  dựa vào phương pháp xấp xỉ bình phương tối thiểu, tức là gọi  $r_1, \dots, r_n$  và  $d_1, \dots, d_n$  là giá trị pixel của  $\mathcal{R}_i$  và  $\mathcal{D}_j$  đã được biến đổi tương ứng, thì:

$$Err = \sum_{i=1}^n (sd_i + o - r_i)^2 \quad (4.2)$$

Để cực tiểu Err, thì đạo hàm theo  $s$  và  $o$  phải =0,

$$Err = n \cdot o^2 + \sum_{i=1}^n (s^2 \cdot d_i^2 + 2 \cdot s \cdot d_i \cdot o - 2 \cdot s \cdot d_i \cdot r_i - 2 \cdot o \cdot r_i + r_i^2) \quad (4.3)$$

$$\frac{\partial Err}{\partial s} = \sum_{i=1}^n (2sd_i^2 + 2d_i - 2od_i) \quad (4.4)$$

$$\frac{\partial Err}{\partial o} = 2no - \sum_{i=1}^n (2sd_i + 2r_i) \quad (4.5)$$

Giải hệ phương trình trên ta được:

$$s = \frac{n \sum_{i=1}^n d_i r_i - \sum_{i=1}^n d_i \sum_{i=1}^n r_i}{n \sum_{i=1}^n d_i^2 - (\sum_{i=1}^n d_i)^2} \quad (4.6)$$

$$o = \frac{1}{n} [\sum_{i=1}^n r_i - s \sum_{i=1}^n d_i] \quad (4.7)$$

Viết lại đơn giản:

$$s = \frac{\alpha}{\beta}; o = \bar{r} - \frac{\alpha}{\beta} \bar{d}; \alpha = \sum_{i=1}^n (d_i - \bar{d})(r_i - \bar{r}); \beta = \sum_{i=1}^n (d_i - \bar{d})^2 \quad (4.8)$$

### 4.3.3. Thuật toán biểu diễn mẫu chất liệu dựa vào Fractal RMBF

- **Vào:** Mẫu chất liệu  $\mathcal{M}$  (một bức ảnh) có kích thước  $m \times m$
- **Ra:** Mô hình  $Model\_Fractal(\mathcal{M})$  (kích thước  $m \times m$ )
- Thuật toán xây dựng mô tả Fractal RMBF gồm các bước sau:

**Bước 1:** Khởi tạo các thông số về kích thước phân hoạch khối dĩa, khối miền:  $k, r$ . Khởi tạo  $Model\_Fractal\mathcal{M} := \{\emptyset\}$

**Bước 2:** Phân hoạch  $\mathcal{M}$  thành tập khối dĩa  $\mathfrak{R} = \{\mathfrak{R}_i\}_{i=1}^N$ , kích thước  $k \times k$ .

**Bước 3:** Phân hoạch  $\mathcal{M}$  thành tập khối miền  $\mathcal{D} = \{\mathcal{D}_j\}_{j=1}^M$ , kích thước  $rk \times rk$

**Bước 4:** Với mỗi khối dĩa  $\mathfrak{R}_i \in \mathfrak{R}, i = 1 \dots N$ , thực hiện:

**Bước 4.1:** Chọn khối miền  $\mathcal{D}_j$  và xác định phép biến đổi  $\mathcal{W}_j$  tương ứng sao cho  $\mathfrak{R}_i \approx \mathcal{W}_j(\mathcal{D}_j)$  như trong phần 4.3.2.

**Bước 4.2:** Tính sai số  $\rho_i = \mathfrak{R}_i - \mathcal{W}_j(\mathcal{D}_j)$ .

**Bước 4.3:** Lưu  $Rec(\mathfrak{R}_i)$  ( vị trí của  $\mathfrak{R}_i$  trên  $\mathcal{M}$  ) và các tham số hình học và ánh sáng ( $a, b, c, d, e, f, s, o$ ) của phép biến đổi  $\mathcal{W}_i$  tương ứng với  $\rho_i$  nhỏ nhất. Tức là  $Model\_Fractal(\mathcal{M}) := Model\_Fractal(\mathcal{M}) \cup \{Rec(\mathfrak{R}_i), \mathcal{W}_i\}$

Như vậy mô hình chất liệu thu được sẽ gồm một tập:

$$Model\_Fractal(\mathcal{M}) = \{\langle Rec(\mathfrak{R}_i), \mathcal{W}_i \rangle\}_{i=1}^K.$$

- Nhận xét:
  - Thuật toán RMBF luôn đảm bảo tính dừng: Việc xác định tập khối dĩa và tập khối miền trong bước 2 và bước 3 là hữu hạn. Ở bước 4, số lần lặp chỉ là N. Do vậy thuật toán luôn đảm bảo tính dừng.

- Thuật toán RMBF luôn đảm bảo tính đúng đắn: Thuật toán sẽ dừng ở bước 4, trong bước 4 ta luôn xác định được  $\rho_i$  nhỏ nhất sao cho  $\rho_i = \mathfrak{R}_i - \mathcal{W}_j(\mathcal{D}_j)$ , lúc đó  $Model\_Fractal(\mathcal{M}) := Model\_Fractal(\mathcal{M}) \cup \{Rec(\mathfrak{R}_i), \mathcal{W}_i\}$  là hoàn toàn đúng đắn theo định nghĩa mô hình Fractal.

- Độ phức tạp tính toán của thuật toán RMBF

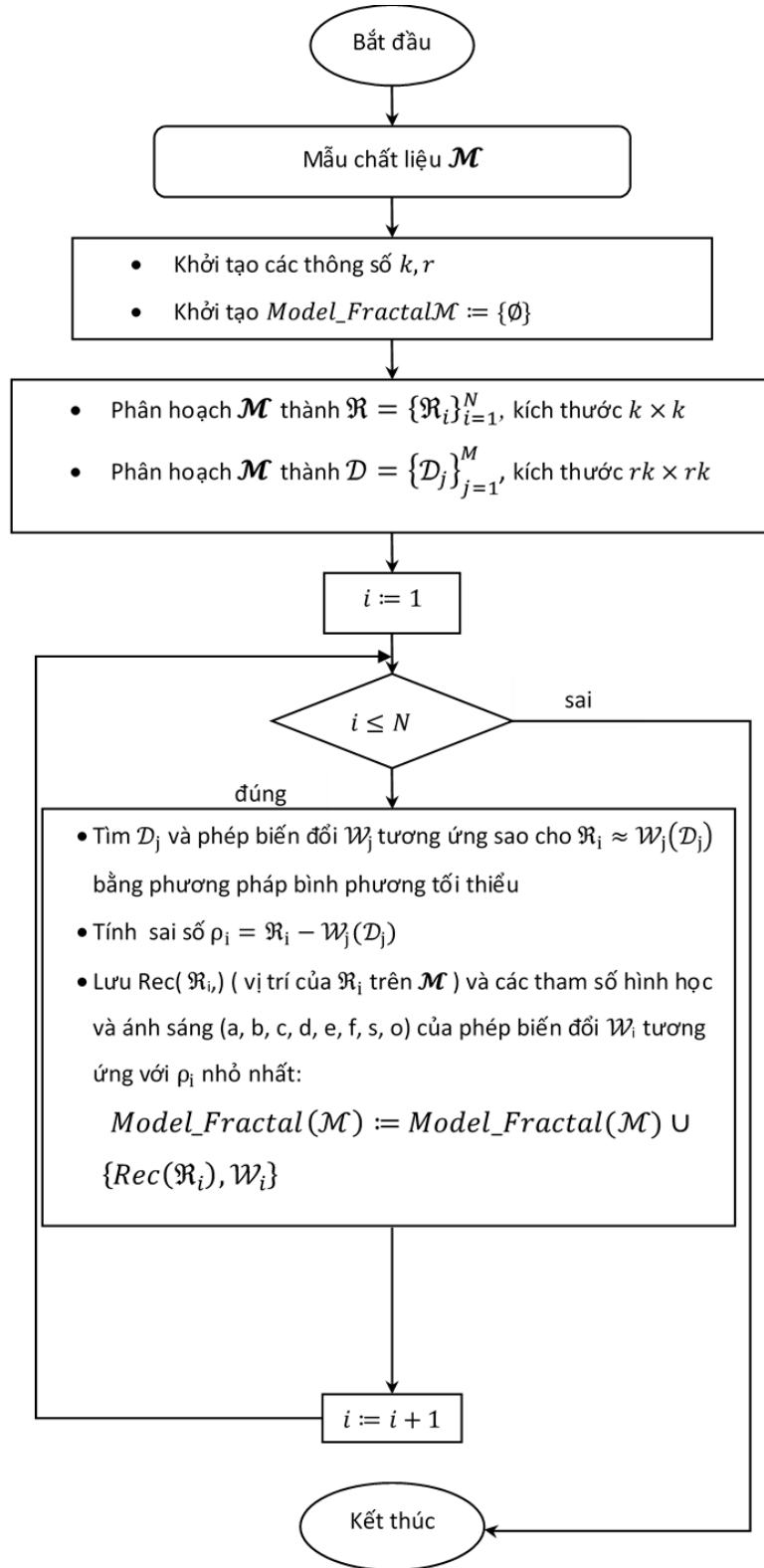
Tính độ phức tạp tính toán theo kích thước mẫu chất liệu, các tham số của các phép tính toán Fractal là cố định.

Phân hoạch chất liệu  $\mathcal{M}$  kích thước  $m \times m$  thành tập các khối dãy không chồng lên nhau  $\mathfrak{R}_i$  kích thước  $k \times k$ , độ phức tạp tính toán sẽ là  $O(\frac{1}{k^2} m^2) \approx O(m^2)$ .

Phân hoạch chất liệu  $\mathcal{M}$  kích thước  $m \times m$  thành tập các khối dãy  $\mathcal{D}_i$  kích thước  $rk \times rk$ , độ phức tạp tính toán sẽ là  $O(m - rk + 1)(m - rk + 1) \approx O(m^2)$ .

Với mỗi khối  $\mathfrak{R}_i$ , ta duyệt tất cả các khối  $\mathcal{D}_j$  và thực hiện phép biến đổi  $\mathcal{W}_j$ , sau đó ta tính sai số  $\mathcal{W}_j(\mathcal{D}_j)$  và lưu lại sai số nhỏ nhất. Ta biết việc tìm phép biến đổi  $\mathcal{W}$  có độ phức tạp  $O(1)$  (hằng số). Ta có độ phức tạp tính toán của quá trình trên là  $O(m^2)O(m^2)(O(1) + O(k^2)) = O(m^2)$ .

- Thuật toán RMBF được biểu diễn dưới dạng sơ đồ khối:



#### 4.4. Thuật toán phát hiện mẫu chất liệu dựa vào Fractal DMBF

Phần này luận án sẽ đề xuất một phương pháp phát hiện mẫu chất liệu trong ảnh dựa vào mô tả fractal của chất liệu, phương pháp này có thể phát hiện hiện quả cho cả hai trường hợp mẫu chất liệu có mô hình và không có mô hình.

Thuật toán DMBF:

- **Vào:** Mẫu chất liệu  $\mathcal{M}$  có kích thước  $m \times m$  được biểu diễn bằng Model\_Fractal ( $\mathcal{M}$ ) và ảnh  $\mathcal{I}$ .
- **Ra:** Tập  $R$  gồm các vùng trong ảnh  $\mathcal{I}$  chứa mẫu chất liệu  $\mathcal{M}$ .
- Thuật toán phát hiện mẫu chất liệu gồm các bước chính sau:

**Bước 1:** Phân hoạch  $\mathcal{I}$  thành tập  $\mathcal{B} = \{\mathcal{B}_k\}_{k=1}^P$ ,  $\mathcal{B}_k$  có kích thước bằng kích thước mẫu chất liệu  $\mathcal{M}$ . Chọn thông số ngưỡng đối sánh  $\varepsilon$  và khởi tạo  $R := \{\emptyset\}$ .

**Bước 2:** Với mỗi khối  $\mathcal{B}_k \in \mathcal{B}, k = 1 \dots P$ , thực hiện:

**Bước 2.1:** Tính  $\mathcal{B}'_k = \cup B'^i_k = \cup \{\mathcal{W}_i(Rec(\mathfrak{R}_i)[\mathcal{B}_k])\}$

**Bước 2.2:** Đối sánh  $\mathcal{B}_k$  với  $\mathcal{B}'_k$ . Nếu  $d(\mathcal{B}_k, \mathcal{B}'_k) < \varepsilon$  ( $\varepsilon$  là ngưỡng khoảng cách xác định trước) thì  $R := R \cup \{Rec(\mathcal{B}_k)\}$ .

- Nhận xét:

Thuật toán DMBF luôn đảm bảo tính dừng và tính đúng đắn: Vì số khối tập  $\mathcal{B}$  là hữu hạn, do đó thuật toán sẽ dừng trong bước 2. Tại bước dừng này, thuật toán xác định được  $d(\mathcal{B}_k, \mathcal{B}'_k) < \varepsilon$  chính xác, khi đó nếu  $d(\mathcal{B}_k, \mathcal{B}'_k) < \varepsilon$  thỏa mãn thì vùng  $\mathcal{B}_k$  chứa mẫu chất liệu với sai số  $\varepsilon$  là hoàn toàn đúng đắn.

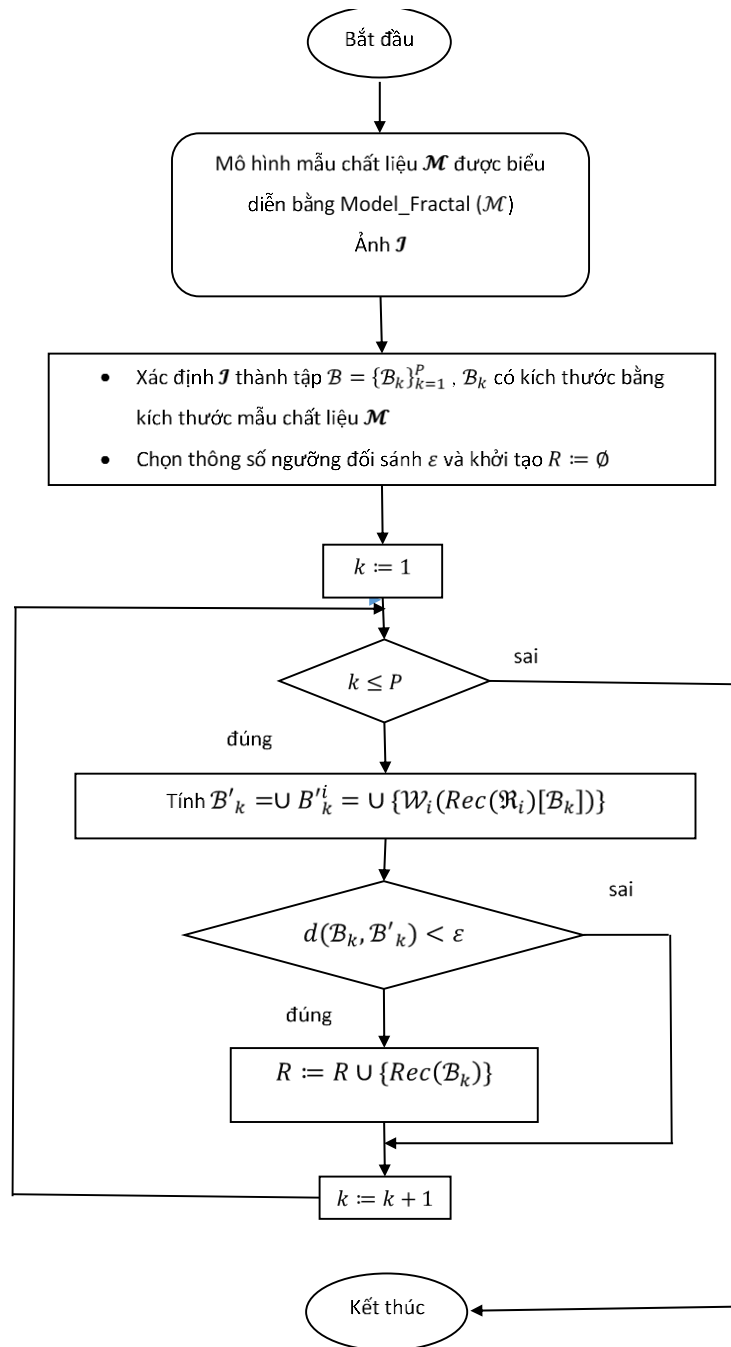
- Độ phức tạp tính toán của thuật toán DMBF

Ta tính độ phức tạp thuật toán theo kích thước ảnh đầu vào, các tham số về mô hình Fractal là cố định. Thuật toán thực hiện bằng cách duyệt mỗi cửa sổ có kích thước bằng kích thước mẫu chất liệu, với mỗi cửa sổ tìm được ta thực hiện phép biến đổi  $\mathcal{W}$  và tính sai số giữa  $\mathcal{W}_j(\mathcal{D}_j)$  và  $\mathfrak{R}_i$ . Giả sử  $n \times n$  là kích thước ảnh vào,  $m \times$



$m$  là kích thước mẫu chất liệu và  $k \times k$  là kích thước của các khối dây con, thì độ phức tạp tính toán của thuật toán sẽ là  $O((n - m + 1)(n - m + 1))(O(1) + O(k^2)) \approx O(n^2)$ .

- Thuật toán DMBF được biểu diễn dưới dạng sơ đồ khối:



#### **4.5. Kết luận chương 4**

Chương này luận án đã đề xuất một kỹ thuật phát hiện mẫu chất liệu trên cơ sở xây dựng mô hình Fractal cho mẫu chất liệu và phương pháp phát hiện mẫu chất liệu dựa vào tiếp cận lý thuyết Fractal. Trên cơ sở ý tưởng lý thuyết hàm lặp IFS của hình học Fractal để tìm ra các đặc trưng ảnh lặp lại như các đường vân, biên hay xương v.v. Đây chính là đặc tính quan trọng của mẫu chất liệu trong thực tế như trong chương 1 đã phân tích. Bước đầu luận án đã nghiên cứu lý thuyết và đưa ra mô hình Fractal cho bài toán phát hiện mẫu chất liệu trên cơ sở đề xuất hai thuật toán biểu diễn và phát hiện mẫu chất liệu nhằm giảm thiểu không gian lưu trữ và thời gian tính toán trong quá trình xây dựng đặc trưng và phát hiện mẫu chất liệu.

## KẾT LUẬN

Phát hiện mẫu chất liệu trong ảnh là một bài toán mở hiện nay của Xử lý ảnh và Thị giác máy do tính đa dạng và phức tạp của các loại chất liệu trong thực tế cũng như nhiều điều kiện thu nhận ảnh trong những tình huống khác nhau. Sự thay đổi ánh sáng, hướng và tỉ lệ do quá trình thu nhận ảnh là những vấn đề hết sức khó khăn trong các nghiên cứu về chất liệu nói chung và bài toán phát hiện mẫu chất liệu từ ảnh nói riêng.

Luận án đã đề xuất nghiên cứu phát hiện mẫu chất liệu trong ảnh theo hướng nghiên cứu các đặc trưng biểu diễn chất liệu được xây dựng từ các đặc trưng ở mức thấp như màu sắc, kết cấu v.v. Các đặc trưng này bất biến với một số phép biến đổi hình học, hoặc ít nhạy với các phép biến đổi nào đó đặc biệt là ánh sáng. Xuất phát từ thực tế đó luận án đã tìm hiểu tổng quan về phát hiện mẫu chất liệu trong ảnh, nghiên cứu các kỹ thuật biểu diễn và phát hiện mẫu chất liệu cùng với ứng dụng giải quyết bài toán phát hiện ảnh số giả mạo trong trường hợp vùng giả mạo bị thay đổi bởi phép quay và phép thay đổi tỉ lệ mà các kỹ thuật khác chưa phát hiện được. Cụ thể, luận án đã đạt được các kết quả chính sau:

- Đề xuất kỹ thuật phát hiện mẫu chất liệu trong ảnh dựa vào đặc trưng bất biến địa phương DMBLIF. Kỹ thuật nhằm giải quyết bài toán đặt ra trong trường hợp mẫu chất liệu cần tìm được cho dưới dạng một ảnh con cho trước. Trên cơ sở kỹ thuật DMBLIF, luận án cũng đề xuất việc ứng dụng kỹ thuật này vào việc giải quyết bài toán phát hiện ảnh giả mạo dạng cắt dán. Kỹ thuật phát hiện ảnh số giả mạo KPFIImage mà luận án đề xuất đã giải quyết được bài toán phát hiện giả mạo dạng cắt dán trong trường hợp có sự thay đổi về tỉ lệ và góc quay của vùng cắt dán.
- Đề xuất một cách biểu diễn chất liệu dựa vào đặc trưng nhiễu nhằm giảm thiểu sự ảnh hưởng của ánh sáng mà các nghiên cứu biểu diễn ảnh rất quan tâm. Đưa ra thuật toán phát hiện mẫu chất liệu DMBNF dựa vào đặc trưng

nhiều, kết quả cho thấy kỹ thuật đề xuất khá hiệu quả trong trường hợp mẫu chất liệu có sự thay đổi ánh sáng lớn.

- Đề xuất kỹ thuật biểu diễn chất liệu bằng hình học Fractal nhằm tìm ra đặc trưng bất biến tỉ lệ toàn cục, đây cũng là vấn đề hiện đang được nhiều nghiên cứu quan tâm. Trên cơ sở đó, luận án cũng đưa ra thuật toán biểu diễn mẫu chất liệu RMBF và thuật toán phát hiện mẫu chất liệu DMBF, thuật toán này giải quyết tốt đối với những mẫu chất liệu có cấu trúc kết cấu cao, đặc biệt cho các loại chất liệu nhân tạo.

#### **Vấn đề có thể nghiên cứu tiếp theo:**

- Nghiên cứu, cải tiến và mở rộng các dạng mẫu chất liệu sang dạng mẫu đối tượng dựa trên việc biểu diễn đối tượng bởi nhiều mẫu chất liệu.
- Nghiên cứu các kỹ thuật nâng cao chất lượng phát hiện mẫu chất liệu, khắc phục các yếu tố về môi trường, đây là vấn đề khó khăn mà các hệ thống xử lý ảnh thường gặp phải.
- Nghiên cứu phát triển các ứng dụng vào các bài toán cụ thể trong thực tiễn.

## **DANH MỤC CÁC CÔNG TRÌNH KHOA HỌC CỦA TÁC GIẢ LIÊN QUAN ĐẾN LUẬN ÁN**

1. Đỗ Năng Toàn, Lê Thị Kim Nga (2007), “Một cách tiếp cận trong phát hiện đối tượng đột nhập”, *Kỷ yếu Hội thảo Quốc gia: Các vấn đề chọn lọc của Công nghệ Thông tin và Truyền thông*, tr. 175-181.
2. Đỗ Năng Toàn, Hà Xuân Trường, Phạm Việt Bình, Lê Thị Kim Nga, Ngô Đức Vĩnh (2008), “Một cải tiến cho thuật toán phát hiện ảnh giả mạo Exact match”, *Kỷ yếu Hội thảo Quốc gia nghiên cứu cơ bản và ứng dụng Công nghệ thông tin – FAIR07*, tr. 161-172.
3. Đỗ Năng Toàn, Lê Thị Kim Nga, Nguyễn Thị Hồng Minh (2010), “Một mô hình nhiễu và ứng dụng trong việc phát hiện chất liệu”, *Tạp chí Khoa học và Công nghệ-Viện Khoa học và Công nghệ Việt Nam* Tập 48 (3), tr. 1-10.
4. Lê Thị Kim Nga, Đỗ Năng Toàn (2010), “Phát hiện ảnh cắt dán giả mạo dựa vào các đặc trưng bất biến”, *Tạp chí Tin học và Điều khiển học- Viện Khoa học và Công nghệ Việt Nam* Tập 26 (2), tr. 185-195.
5. Lê Thị Kim Nga (2010), “Phát hiện chất liệu, tiếp cận và ứng dụng”, *Tạp chí Khoa học và Công Nghệ - Đại học Thái Nguyên* Tập 69 (7), tr. 25-31.
6. Lê Thị Kim Nga, Đỗ Năng Toàn (2010), “Một cách tiếp cận cho phát hiện chất liệu ảnh”, *Kỷ yếu Hội thảo Quốc gia: Các vấn đề chọn lọc của Công nghệ Thông tin và Truyền Thông*, tr. 202-213.
7. Lê Thị Kim Nga, Đinh Mạnh Tường (2010), “Phát hiện chất liệu dựa vào nhiễu”, *Kỷ yếu Hội thảo Quốc gia nghiên cứu cơ bản và ứng dụng Công nghệ thông tin – FAIR09*, tr. 207-214.
8. Do Nang Toan, Le Thi Kim Nga (2011), “Materials Detection Based on Fractal Approach”, *ACM Proceedings of the 9th International Conference on Advances in Mobile Computing & Multimedia (MoMM2011)*, pp. 281-284.

## TÀI LIỆU THAM KHẢO

1. Adelson E.H. (2001), “On Seeing Stuff: The Perception of Materials by Humans and Machines”, *In Proceedings of the SPIE* Vol. 4299, pp. 1–12.
2. Arai K., Bu X.Q. (2007), “ISODATA Clustering with Parameter (Threshold for Merge and Split) Estimation based on Genetic Algorithm”, *Reports of the Faculty of Science and Engineering* Vol. 6(1), pp. 17-23.
3. Asada H., Brady M. (1986), “The Curvature Primal Sketch”, *Pattern Analysis and Applications* Vol. 8(1), pp. 2-14.
4. Bar-Hillel A., Hertz T., Shental N., Weinshall D. (2003), “Learning Distance Functions using Equivalence Relations”, *Proceedings of the Twentieth International Conference on Machine Learning (ICML-2003)* Vol. 20, pp. 11-18.
5. Barnsley M. (1988), *Fractals Everywhere*, Academic Press, San Diego - USA.
6. Beaudet P.R (1987), “Rotationally Invariant Image Operators”, *International Joint Conference on Pattern Recognition*, pp. 579-583
7. Belongie S., Malik J., Puzicha J. (2002), “Shape Matching and Object Recognition using Shape Contexts”, *IEEE Transactions Pattern Analysis and Machine Intelligence* Vol. 24 (4), pp. 509-522.
8. Bergen J.R., Adelson E.H. (1988), “Early Vision and Texture Perception”, *Reprinted From Nature* Vol. 333 (6171), pp. 363-364.
9. Bi J., Bennett K., Embrechts M., Breneman M., Song M. (2003), “Dimensionality Reduction via Sparse Support Vector Machines”, *Journal Machine Learning Research* Vol. 3, pp. 1229-1243.
10. Bileschi S., Wolf L. (2005), “A Unified System for Object Detection, Texture Recognition, and Context Analysis based on The Standard Model Feature Set”, *Proceeding In British Machine Vision Conference*, pp. 175-185.

11. Boiman O., Shechtman E., Irani M. (2008), "In Defense of Nearest-neighbor based Image Classification", *In Computer Vision and Pattern Recognition*, pp. 1-8.
12. Bo L., Sminchisescu C. (2009), "Efficient Match Kernels Between Sets of Features for Visual Recognition", *Proceeding In Neural Information Processing Systems*, pp. 1-9.
13. Boyer C.B., Merzbach (1989), *A History of Mathematics*, John Wiley & Sons, 2<sup>nd</sup> Edition, New York.
14. Boykov Y., Pjolly M. (2004), "Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D images", *IEEE International Conference on Computer Vision and Pattern Recognition* Vol. 2, pp. 731-738.
15. Bo L., Ren X., Fox D. (2010), "Kernel Descriptors for Visual Recognition", *In Neural Information Processing Systems*, pp. 1-9.
16. Broadhurst R.E. (2005), "Statistical Estimation of Histogram Variation for Texture Classification", *Proceeding of Fourth International Workshop Texture Analysis and Synthesis*, pp. 25-30.
17. Brown M., Lowe D.G. (2002), "Invariant Features from Interest Point Groups", *British Machine Vision Conference (BMVC 2002)*, pp. 656-665.
18. Buades A., Coll B., Morel J.M. (2005), "A Non-Local Algorithm for Image Denoising", *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition* Vol. 2, pp. 60-65.
19. Caputo B., Hayman E., Mallikarjuna P. (2005), "Class-Specific Material Categorisation", *Proceeding of 10th International Conference on Computer Vision* Vol. 2, pp. 1597-1604.
20. Caputo B., Hayman E., Fritz M., Eklundh J.O. (2010), "Classifying Materials in The Real World", *Image and Vision Computing* Vol. 28 (1), pp. 150–163.
21. Chen J.L., Kundu A. (1994), "Rotation and Gray Scale Transform Invariant Texture Identification using wavelet Decomposition and Hidden Markov

- Model”, *IEEE Transaction Pattern Analysis and Machine Intellegent* Vol. 16 (2), pp. 208-214.
22. Cohen F.S., Fan Z., Patel M.A.S. (1991), “Classification of Rotated and Scaled Textured Images using Gaussian Markov Field Models”, *IEEE Transactions Pattern Analysis and Machine Intelligence* Vol. 13 (2), pp. 192–202.
  23. Cula O.G., Dana K.J. (2001), “Compact Representation of Bidirectional Texture Functions”, *In Proceeding of Computer Vision and Pattern Recognition* Vol. 1, pp. 1041-1047.
  24. Cula O.G., Dana K.J. (2004), “3D Texture Recognition Using Bidirectional Feature Histograms”, *Journal of Computer Vision* Vol. 59 (1), pp. 33-60.
  25. Dana K.J., Ginneken B.V., Nayar S.K., Koendrink J.A.N.J. (1999), “Reflectance and Texture of Real – World Surface”, *Association for Computing Machine Transactions on Graphics* Vol. 18 (1), pp. 1-34.
  26. Debevec P., Hawkin T., Tchou C., Duiker H.P., Sagar M. (2000), “Acquiring the Reflectance field of a Human face”, *In Association for Computing Machinery's Special Interest Group on Computer Graphics and Interactive Techniques (ACM SIGGRAPH)*, pp. 145-156.
  27. Diane H., Bo L., Ren X. (2012), “Toward Robust Material Recognition for Everyday Objects”, *Everyday Material Recognition*, pp. 1-11.
  28. Do Nang Toan, Le Thi Kim Nga, Ha Xuan Truong (2007) , “Một cải tiến cho thuật toán phát hiện ảnh giả mạo Exact match”, *Kỷ yếu Hội thảo FAIR'2007*, tr.161-172.
  29. Dorko G., Schmid C. (2003), “Selection of Scale-Invariant Parts for Object Class Recognition”, *Proceeding of Ninth International Conference on Computer Vision*, pp. 634-640.
  30. Dror R., Adelson E.H., Willsky A.S. (2001), “ Recognition of Surface Reflectance Properties from A Single Image under Unknown Real World Illumination”, *In IEEE Workshop on Identifying Object across Variation in Lighting*, pp. 1-8.



31. Duda R.O., Hart P.E., Stork D.G. (2001), *Pattern Classification*, seconded, John Wiley & Sons, New York.
32. Ebrahimi M., Vrscay E.R. (2008), "Examining The Role of Scale in The Context of The non-local-means filter", *In Image Analysis and Recognition* Vol. 5112, pp. 170-181.
33. Ebrahimi M., Vrscay E.R. (2008), "Multiframe Super Resolution with no Explicit Motion Estimation", *In Proceedings of The 2008 International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV 2008, Las Vegas, Nevada, USA)*, pp. 1-7.
34. Ebrahimi M., Vrscay E.R. (2003), *Self-Similarity in Imaging, 20 Year After: Fractals Everywhere, Fractal Geometry*, John Wiley & Sons, New York.
35. Fei-Fei L., Fergus R., Perona P. (2007), "Learning Generative Visual Models from Few Training Examples: An Incremental Bayesian Approach Tested on 101 Object Categories", *Computer Vision and Image Understanding* Vol. 106 (1), pp. 59-70.
36. Fergus R., Perona P., Zisserman A. (2003), "Object Class Recognition by Unsupervised Scale-Invariant Learning", *In Proceedings of IEEE Computer Vision and Pattern Recognition* Vol. 2, pp. 264-271.
37. Ferrari V., Tuytelaars T., Gool L.V. (2004), "Simultaneous Object Recognition and Segmentation by Image Exploration," *Proceeding of Eighth European Conference on Computer Vision*, pp. 40-54.
38. Fisher Y. (2003), *Fractal Image Compression*, Springer Verlag, New York.
39. Fowlkes C., Martin D., Malik J. (2003), "Learning Affinity Functions for Image Segmentation: Combining Patch-Based and Gradient-Based Approaches", *Proc. IEEE Conf. Computer Vision and Pattern Recognition* Vol. 2, pp. 54-61.
40. Fridrich J., Soukal D., Lukáš J. (2003), "Detection of Copy-Move Forgery in Digital Images", *Proceeding of Digital Forensic Research Workshop*, pp. 1-10.

41. Fridrich J. (1999), "Methods for Tamper Detection in Digital Images", *Proceeding of Association for Computing Machine Workshop on Multimedia and Security*, pp. 19-23.
42. Harris C., Stephens M. (1988), "A Combined Corner and Edge Detector", *In Proceedings of the 4th Alvey Vision Conference*, pp. 147-151.
43. Hauage D.C., Snavely N. (2012), "Image Matching using Local Symmetry Features", *Proceeding of IEEE*, pp. 206 - 213.
44. Hayman E., Caputo B., Fritz M., Eklundh J.O. (2004), "On the Significance of Real-World Conditions for Material Classification", *Proceeding of Eighth European Conference on Computer Vision Vol. 4*, pp. 253-266.
45. Hays J., Leordeanu M., Efros A., Liu Y. (2006), "Discovering Texture Regularity as a Higher-Order Correspondence Problem", *Proceeding of Ninth European Conference on Computer Vision Vol. 2*, pp. 522-535.
46. Hutchinson J. (1981), "Fractals and Self Similarity", *Indiance University Mathematics Journal Vol. 30 (5)*, pp. 713-747.
47. Jacquin A.E. (1989), *A Fractal Theory of Iterated Markov Operators with Applications to Digital Image Coding*, Ph.D Thesis, Georgia Institute of Technology.
48. Keys E.G. (1981), "Cubic Convolution Interpolation for Digital Image Processing", *IEEE Transactions Acoustics, Speech, and Signal Processing Vol. 29(6)*, pp. 1153-1160.
49. Khriji L., El-Metwally K. (2006), "Rational-Based Particle Swarm Optimization for Digital Image Interpolation", *International Journal of Intelligent Technology Vol. 1 (3)*, pp. 515-519.
50. Leibe B., Schiele B. (2003), "Interleaved Object Categorization and Segmentation", *Proceeding of 14th British Machine Vision Conference*, pp. 759-768.

51. Liu C., Sharan L., Adelson E.H., Rosenholtz R. (2010), “Exploring Features in a Bayesian Framework for Material Recognition”, *In Computer Vision and Pattern Recognition*, pp. 1-8.
52. Jan-Mark G. (2001), “Color Invariance”, *IEEE Transactions of Pattern Analysis and Machine Intelligence* Vol. 23 (12), pp. 1338-1350.
53. Jian G.W., Li J., Lee C.Y., Yau W.Y. (2010), “Dense SIFT and Gabor Descriptor-based Face Representation with Application to Gender Recognition”, *Control Automatic Robotics & Vision 2010*, pp. 1860 -1864.
54. Henry K.M., Ponce J. (2006), “A Geodesic Active Contour Framework for Finding Glass”, *Proceeding of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1-8.
55. Lazebnik S., Schmid C., Ponce J. (2006), “Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories”, *In Computer Vision and Pattern Recognition*, pp. 2168–2178.
56. Leung T., Malik J. (2001), “Representing and Recognizing The Visual Appearance of Materials using Three-dimensional Textons”, *Journal of Computer Vision* Vol. 43 (1), pp. 29–44.
57. Lee S., Shamama D.A., Gooch B. (2006), “Detecting False Captioning using Common-Sense Reasoning”, *Digital Investigation, DFRWS Published by Elsevier Ltd*, pp. 65-70.
58. Liebelt J., Schmid C., Schertler K. (2008), “Viewpoint-Independent Object Class Detection using 3d Feature Maps”, *In proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2008*, pp. 1-8
59. Liu X., Li R. (2012), “An Iris Recognition Approach with SIFT Descriptors”, *Advanced Intelligent Computing Theories and Application with Aspects of Artificial Intelligence Lecture Note in Computer Science 2012* Vol. 6839, pp. 427-434.
60. Lowe D.G. (2004), “Distinctive Image Features from Scale-invariant Keypoints”, *Journal of Computer Vision* Vol. 60 (2), pp.91-110.

61. Lowe D.G. (1988), "Organization of Smooth Image Curves at Multiple Scales", *International Conference on Computer Vision*, pp. 558–567.
62. Luka's J. (2001), "Digital Images Authentication Using Image Fitting Techniques", *Proceedings of ALGORITMY 2000 Conference on Scientific Computing*, pp. 236-244.
63. Lazebnik S., Schmid C., Ponce J. (2003), "Sparse Texture Representation Using Affine-Invariant Neighborhoods", *Proceeding of Conference on Computer Vision and Pattern Recognition*, pp. 319-324.
64. Lazebnik S., Schmid C., Ponce J. (2005), "A Sparse Texture Representation Using Local Affine Regions", *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 27 ( 8), pp. 1265-1278.
65. Mahdian B., Stanislav S. (2007), "Detection of Copy-Move Forgery using A Method based on Blur Moment Invariant", *Forensic Science International* Vol. 171, pp. 180-189.
66. Madelbrot B. (1975), *Les Object Fractals: Forme, Hasard Diel Dimension*, Paris: Flammarion.
67. Mao J., A.K. Jain A.K. (1999), "Texture Classification and Segmentation using Multiresolution Simultaneous Autoregressive Models", *Pattern Recognition* Vol. 25, pp. 173-188.
68. Mihcak M.K., Kozintsev I. (1999) "Spatially Adaptive Statistical Modeling of Wavelet Image Coefficients and its Application to Denoising" , *Proceeding of IEEE Int. Conference on Acoustics, Speech, and Signal Processing, Phoenix, Arizona* Vol. 6, pp. 3253–3256.
69. Mikolajczyk K., Schmid C. (2001), "Indexing based on Scale Invariant Interest Point", *In Proceedings of International Conference on Computer Vision* Vol. 1, pp. 525-531.
70. Mikolajczyk K., Schmid C. (2005), "A Performance Evaluation of Local Descriptors", *IEEE Transaction on Pattern Analysis and Machine Intelligent* Vol. 27 (10), pp. 1615-1630.

71. Micah K.J., Farid H. (2006), “Exposing Digital Forgery Through Chromatic Aberration”, *ACM Multimedia and Security Workshop*, pp. 1-8.
72. Monro D.M., Dudbridge F. (1995), “Rendering Algorithms for Deterministic Fractal”, *IEEE Computer Graphics and Applications* Vol. 15 (1), pp. 32-41.
73. Myna A.N., Venkateshmurthy M.G., Patil C.G. (2007), “Detection of Region Duplication Forgery In Digital Images Using Wavelet and Log-Polar Mapping”, *International Conference on Computational Intelligence and Multimedia Applications (IEEE 2007)*, pp. 371-377.
74. Nadia M., Al- Saidi G., Rush M., Sard W.W., Ainun M. O. (2012), “Password Authentication Based on Fractal Coding Scheme”, *Hindawi Publishing Corporation Journal of Applied Mathematics*, pp. 1-18.
75. Nadia M., Al-Saidi G. (2012), “An Efficient Signcryption Method using Fractal Image Coding Scheme”, *International Journal of Applied Mathematics and Information* Vol. 6, pp. 189–197.
76. Narasimhan S.G., Ramesh V., Nayar S.K. (2003), “A Class of Photometric Invariants: Separating Material from Shape and Illumination”, *9<sup>th</sup> IEEE Transactions International Conference on Computer Vision*, pp. 1-8.
77. Ojala T., Pietikainen M., Maenpaa T. (2002), “Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns”, *IEEE Transactions Pattern Analysis and Machine Intelligence* Vol. 24 (7), pp. 971-987.
78. Opelt A., Fussenegger M., Pinz A., Auer P. (2004), “Weak Hypotheses and Boosting for Generic Object Detection and Recognition”, *Proceeding of Eighth European Conference on Computer Vision*, pp. 71-84.
79. Perko R., Bischof H. (2004), “Efficient Implementation of Higher Order Image Interpolation”, *Proceeding Of the International Conference in Central Europe on Computer Graphic, Visualization and Computer Vision*, pp. 213-218.
80. Pont S.C., Koendrink J.J. (2005). “ Bidirectional Texture Contrast Function”, *Journal of Computer Vision* Vol. 62 (1), pp. 17-34.

81. Roland W. F., Heinrich H. B. (2005), “Low-Level Image Cues in the Perception of Translucent Materials”, *ACM Transactions on Applied Perception* Vol 2 (3), pp. 346-382.
82. Rothganger F., Lazebnik S., Schmid C., Ponce J. (2006), “3D Object Modeling and Recognition Using Local Affine-Invariant Image Descriptors and Multi - View Spatial Constraints”, *International Journal of Computer Vision* Vol. 66 (3), pp. 231-259.
83. Rubner Y., Tomasi C., Guibas L. (2000), “The Earth Mover’s Distance as A Metric for Image Retrieval”, *International Journal of Computer Vision* Vol. 40 (2), pp. 99–121.
84. Schaffalitzky F., Zisserman A. (2001), “Viewpoint invariant texture matching and wide baseline stereo”, *In Proceeding International Conference on Computer Vision* Vol. 2, pp. 636–643.
85. Schmid C., Mohr R. (1997), “Local Grayvalue Invariants for Image Retrieval”, *IEEE Transactions Pattern Analysis and Machine Intelligence* Vol. 19 (5), pp. 530-534.
86. Schmid C. (2001), “Constructing Models for Content-Based Image Retrieval”, *Proceeding IEEE Conference on Computer Vision and Pattern Recognition* Vol. 2, pp. 39-45.
87. Se S., Lowe D.G., Little J. (2002), “Global Localization Using Distinctive Visual Features”, *Proceeding of International Conference on Intelligent Robots and Systems*, pp. 226-231.
88. Sharan B.L., Rosenholtz R., Adelson E.H. (2009). “Material perception: What can you see in a brief glance?”, *Journal of Vision* Vol. 9 (8), pp. 784-794.
89. Sharan B.L., Liu C., Rosenholtz R., Adelson E.H. (2013), “Recognizing Materials Using Perceptually Inspired Features”, *International Journal of Computer Vision* Vol. 103, pp. 348-371.
90. Takagi M., Fujiyoshi H. (2007), “Road Sign Recognition using SIFT feature”, *Symposium on Sensing via Image Information*, pp. 1-9.

91. Takayasu H. (1990), *Fractal in the Physical Science*, Manchester University Press.
92. Teynor A. (2009), *Visual Object Class Recognition using Local Descriptions*, Ph.D Thesis, Freiburg University.
93. Teynor A., Burkhardt H. (2007), “Fast codebook generation by sequential data analysis for object classification”, *In Proceedings of the 3<sup>rd</sup> International Symposium on Visual Computing*, pp. 1-8.
94. Teynor A., Burkhardt H. (2008), “Wavelet – based Salient Points with Scale Information for Classification”, *In IEEE 2008*, pp. 1-4.
95. Teynor A., Burkhardt H. (2007), “Patch Based Localization of Visual Object Class Instance”, *MVA2007 IAPR Conference on Machine Vision Applications*, pp. 211-214.
96. Teynor A., Kowarschik W. (2005), “Compressed Domain Image Retrieval using JPEG2000 and Gaussian Mixture Models”, *In 8th International Conference on Visual Information Systems*, pp. 1-8.
97. Thomas A., Ferrari V., Leibe B., Tuytelaars A.T., (2006), “Towards Multi-View Object Class Detection”, *In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1-8.
98. Thomas A., Ferrari V., Leibe B., Tuytelaars A.T., Gool L.V. (2009), “Using Multi-View Recognition and Meta-data Annotation to Guide a Robot’s Attention”, *The International Journal of Robotics Research* Vol. 28 (8), pp. 976-998.
99. Tuytelaars A.T., Gool L.V. (2004), “Matching Widely Separated Views based on Affinely Invariant Neighbourhoods”, *International Journal of Computer Vision* Vol. 59 (1), pp. 61-85.
100. Tuytelaar A.T., Mikolajczyk K. (2007), “Local Invariant Feature Detectors: A Survey”, *Computer Graphic and Vision* Vol. 3 (3), pp. 177-280.

101. Vacha P., Haindl M. (2010). "Illumination Invariants Based on Markov Random Fields", *Pattern Recognition Recent Advances, Adam Herout*, pp. 255-272.
102. Vacha P., Haindl M. (2010), "Natural Material Recognition with Illumination Invariant Textural Features", *In IEEE of International Conference on Pattern Recognition*, pp. 858-861.
103. Varma M., Ray D. (2007), "Learning The Discriminative Power-Invariance Trade-Off", *In International Conference on Computer Vision*, pp. 1-8.
104. Varma M. (2004), *Statistical Approaches to Texture Classification*, Ph.D Thesis, University of Oxford.
105. Varma M., Garg R. (2007), "Locally Invariant Fractal Features for Statistical Texture Classification", *Proceeding of 11th Intelligence Conference on Computer Vision*, pp. 1-8.
106. Varma M., Zisserman A. (2003), "Texture Classification: Are Filter Banks Necessary?", *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition* Vol. 2, pp. 691-698.
107. Varma M., Zisserman A. (2003), "Classifying Images of Materials: Achieving Viewpoint and Illumination Independence", *Computer Vision and Pattern Recognition* Vol. 3, pp. 255-271.
108. Varma M., Zisserman A. (2004), "Unifying Statistical Texture Classification Frameworks", *Image and Vision Computing* Vol. 22 (14), pp. 1175-1183.
109. Varma M., Zisserman A. (2009), "A Statistical Approach to Material Classification using Image Patch Exemplars", *IEEE Transactions Pattern Analysis and Machine Intelligence* Vol. 31 (11), pp. 2032–2047.
110. Weinberger K.Q., Saul L.K. (2009), "Distance metric learning for large margin nearest neighbor classification", *The Journal of Machine Learning Research* Vol. 10, pp. 207–244.
111. Welsted S. (1999), *Fractal and wavelet image compression techniques*, SPIE Press.



112. Winder S., Brown M. (2007), “Discriminative Learning of Local Image Descriptors”, *In Computer Vision and Pattern Recognition* Vol. 33, pp. 43-57.
113. Winder S., Hua G., Brown M. (2009), “Picking the best Daisy”, *In Computer Vision and Pattern Recognition*, pp.1-8.
114. Wu C., Frahm J.M., Pollefeys M. (2010), “Detecting Large Repetitive Structures with Salient Boundaries”, *European Conference on Computer Vision*, pp. 142-155.
115. Xia Y., Feng D., Zhao R.C., Zhang Y. (2010), “Multifractal Signature Estimation for Textured Image”, *Segmentation and Pattern Recognition Letter*, pp. 163-169.
116. Xu Y., Huang S., Ji H., Fermuller C. (2012), “Scale Space Texture Description on SIFT – Like textons”, *Computer Vision and Image Understanding, Elsevier 2012* Vol. 116, pp. 999-1013.
117. Yaowei W., Weigiang W., Yanfei W. (2004), “ A Region based Image Matching Method with Regularied SAR Model”, *PCM’04 Proceeding of the 5<sup>th</sup> Pacific Rim conference on Advances in Multimedia Information Processing*, pp. 263-270.
118. Zhang J., Lazebnik S., Schmid C. (2007), “Local Features and Kernels for Classification of Texture and Object Categories: a Comprehensive Study”, *International Journal of Computer Vision* Vol. 73 (2), pp. 213-238.
119. Zhao J., Cao Z., Zhou M. (2007), “ SAR Image Denoising based on Wavelet-Fractal Analysis”, *System Engineer Electronic* Vol. 18, pp. 45-48.

## PHỤ LỤC

- A1. Thuật toán phát hiện mẫu chất liệu dựa vào đặc trưng bất biến địa phương DMBLIF

Algorithm DMBLIF (M, I);

{

    Init();

    Mc=(M.with/2, M.heigh/2);

    R=∅;

    Do {  $I_c = \emptyset$ ;

        For each  $F_M^i \in FM$

$\{F_I^{p'} = nearest(FI, F_M^i);$

            If ( $F_I^{p'} = null$ ) continue;

$\Delta_i = vector\_dinh\_vi(F_M^i, Mc);$

$I_c^i = center(F_I^{p'}, \Delta_i)$

            If ( $I_c^i = null$ ) continue;

$I_c = I_c \cup \{I_c^i\}$

        }

    If  $size(I_c) \leq \gamma$  break;

    C=ISODATA( $I_c, \epsilon$ );

    b=false;

    for each  $C^i \in C$

        { if  $size(C^i) \geq \gamma$

```

        { FC =  $\emptyset$ ;
        For each  $\langle I_c^j, F_l^j \rangle \in \langle C^i, FI \rangle$ 
        { FC = FC  $\cup$   $\{F_l^j\}$ 
        }
        R=R $\cup$  Rec(FC);
        FI=FI\FC;
        B=true;
        }
    }
    If b=false
    { break;
    }
} While (FI  $\neq$   $\emptyset$ ) return R;
}

```

- A2. Thuật toán phát hiện ảnh số giả mạo KPFIImage

```

KPFIImage(I)
{ Init();
  RF= $\emptyset$ ;
  For each  $X^i \in I$  and size( $X^i$ )= (B, B)
  {  $X_{p,q}^i = f(X^i, p, q)$ ;
    R=DMBLIF( $X_{p,q}^i, I$ )
    If size(R)  $\geq$  2
    { RF=RF $\cup$  R
    }
  }
}

```

```

    }
    Return RF;
    }

```

- A3. Thuật toán biểu diễn mẫu chất liệu dựa vào nhiễu RMBN

Algorithm RMBN(M)

```

{
     $n = size(M); return \frac{1}{n} \sum_{i=1}^n noise(M_i)$ 
}

```

- A4. Thuật toán phát hiện mẫu chất liệu dựa vào đặc trưng nhiễu DMBNF

Algorithm DMBNF (M,I)

```

{ Init (); R=∅; I' = noise(I);
    For each  $R_i \in I'$  and  $size(RI') = size(M)$ 
        {  $\rho_{Ri} = correlate(R_i, M); p_{Ri} = 1 - G(\rho_{Ri});$ 
            If ( $p_{Ri} < \alpha$ )
                {
                    R=R∪ Rec( $R_i$ )
                }
        }
    }
    Return R
}

```

- A5. Thuật toán biểu diễn chất liệu dựa vào hình học Fractal RMBF

Algorithm RMBF (M)

```

{ Init=();
  MF=∅;
  R=split(M,k);
  For each  $R_i \in R$ 
    { for each  $D_j \in M$  and  $size(D_j) = (rk, rk)$ 
      {  $\mathcal{W}_j = calculate \mathcal{W}(R_i, D_j)$ ;
         $\rho_i = \|R_i - \mathcal{W}_j(D_j)\|$ 
        If ( $\rho_i < \rho$ )
          {  $\rho = \rho_i$ ;  $\mathcal{W} = \mathcal{W}_j$  }
        }
      MF=MF  $\cup$  {Rec( $R_i$ ),  $\mathcal{W}$ }
    }
  Return MF
}

```

- A6. Thuật toán phát hiện mẫu chất liệu dựa vào hình học Fractal DMBF

Algorithm DMBF (M, I)

```

{ Init();
  R = ∅;
  For each  $B_k \in I$  and  $size(B_k)=size(M)$ 
    {  $B'_k = \cup B_k^i = \cup \{\mathcal{W}_i(Rec(R_i)[B_k])\}$ ;

```

```
If ( $d(B_k, B'_k) < \varepsilon$ )
    {
         $R = R \cup Rec(B_k)$ ;
    }
}
Return R
}
```