**ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**

**KHOA CÔNG NGHỆ THÔNG TIN**

**🙠🕮🙢**

**Đồ án 9**

**Khảo sát phép biến đổi Wavelet  
 trên ảnh RGB-D**

GVHD: PGS. TS. Lý Quốc Ngọc

Nhóm thực hiện:

Ngành Khoa Học Máy Tính – Cao học khóa 23

1. Hồ Quang Minh - 1211042
2. Đỗ Đặng Minh - 1311015
3. Huỳnh Công Toàn - 1311026
4. Dương Xuân Long - 1311048
5. Hồ Văn Tấn - 1311058

*Tp. Hồ Chí Minh, tháng 10 năm 2014*

# Thông tin nhóm

|  |  |  |  |
| --- | --- | --- | --- |
| **MSHV** | **Họ tên** | **Email** | **Điện thoại** |
| 1211042 | Hồ Quang Minh | [minhho242@gmail.com](mailto:minhho242@gmail.com) | 093-332-1322 |
| 1311015 | Đỗ Đặng Minh (nhóm trưởng) | [masterminh219@gmail.com](mailto:masterminh219@gmail.com) | 0168-993-5242 |
| 1311026 | Huỳnh Công Toàn | [alex7huynh@gmail.com](mailto:alex7huynh@gmail.com) | 0121-516-1090 |
| 1311048 | Dương Xuân Long | [xuanlong.8888@gmail.com](mailto:xuanlong.8888@gmail.com) | 097-357-0042 |
| 1311058 | Hồ Văn Tấn | [tanhv90@gmail.com](mailto:tanhv90@gmail.com) | 090-290-9334 |

# Danh mục các kí hiệu, chữ viết tắt và ý nghĩa

|  |  |  |
| --- | --- | --- |
| **Từ viết tắt** | **Nghĩa tiếng Anh** | **Nghĩa tiếng Việt** |
| DWT | Discrete wavelet transform | Biến đổi wavelet rời rạc |
| CWT | Continuous wavelet transform | Biến đổi wavelet liên tục |
| FWT | Fast wavelet transform | Biến đổi wavelet nhanh |
| SPIHT | Set partitioning in hierarchical tree | Mã hóa phân cấp theo vùng |
| EZW | Embedded Zerotree Wavelet | Wavelet cây zero |
| DoF | Depth of field | Độ sâu trường ảnh |

Mục lục

[Thông tin nhóm 1](#_Toc397351428)

[Danh mục các kí hiệu, chữ viết tắt và ý nghĩa 1](#_Toc397351429)

[I. Giới thiệu 3](#_Toc397351430)

[II. Wavelet trên ảnh màu RGB 3](#_Toc397351431)

[1) Biến đổi wavelet rời rạc 3](#_Toc397351432)

[2) Biến đổi wavelet liên tục 3](#_Toc397351433)

[III. Wavelet trên ảnh độ sâu 3](#_Toc397351434)

[1) Ảnh độ sâu 3](#_Toc397351435)

[2) Ảnh màu-độ sâu 3](#_Toc397351436)

[IV. Ứng dụng của Wavelet 3](#_Toc397351437)

[1) Nén ảnh JPEG 2000 3](#_Toc397351438)

[2) Nén video 6](#_Toc397351439)

[3) Nén âm thanh 6](#_Toc397351440)

[4) Lưu trữ vân tay điện tử 6](#_Toc397351441)

[5) Chứng thực vân tay 7](#_Toc397351442)

[6) Giảm nhiễu 7](#_Toc397351443)

[7) Các ứng dụng khác 7](#_Toc397351444)

[V. Chương trình minh họa 8](#_Toc397351445)

[1) Tập dữ liệu ảnh 8](#_Toc397351446)

[2) Môi trường lập trình 8](#_Toc397351447)

[3) Mô tả ứng dụng 8](#_Toc397351448)

[VI. Tổng kết 8](#_Toc397351449)

[VII. Tài liệu tham khảo 8](#_Toc397351450)

# Giới thiệu

# Wavelet trên ảnh màu RGB

1. Biến đổi wavelet liên tục

Biến đổi Wavelet lien tục (Continuous Wavelet Transform – CWT) của một hàm f(t) được bắt đầu từ một hàm Wavelet mẹ . Hàm wavelet mẹ có thể là bất kỳ hàm số thực hoặc một hàm phức có thể thỏa mãn các tính chất sau:

* Tích phân suy rộng trên toàn bộ trục t của hàm là bằng 0:



* Tích phân năng lược của hàm trên toàn bộ trục t là một số hữu hạn; (hay còn gọi là bình phương khả tích):



Ở đây, bình phương khả tích nghĩa là tích phân của bình phương của [giá trị tuyệt đối](http://vi.wikipedia.org/wiki/Gi%C3%A1_tr%E1%BB%8B_tuy%E1%BB%87t_%C4%91%E1%BB%91i) của hàm số đó là hữu hạn. *Modulo bằng nhau hầu như khắp nơi* nghĩa là các hàm số này có thể được xác định nếu và chỉ nếu chúng bằng nhau *bên ngoài một tập có độ đo bằng 0*. Như vậy, hàm wavelet này sẽ thuộc không gian L2(R) các hàm bình phương khả tích. Sau khi hàm Wavelet được lựa chọn, biến đổi Wavelet liên tục của một hàm bình phương khả tích f(t) được tính theo công thức như sau:

Trong đó:

quyết định vị trí của wavelet trên miền thời gian.

quyết định vị trí của wavelet trên miền tần số.

là wavelet mẹ (Daubechies-n, Mexican hat, Morel,…).

Biến đổi này là một hàm của hai tham số thực và . Đấu \* ký hiệu là liên hợp phức của . Nếu chúng ta định nghĩa hàm wavelet con theo biểu thức sau:

Như vậy, hàm Wavelet liên tục của hàm f(t) có thể được biểu diễn như sau:

Với công thức như trên, hàm Wavelet liên tục (CWT) là tích vô hướng của hai hàm và . Giá trị là hệ số chuẩn hóa để đảm bảo rằng tích phân năng lượng của hàm độc lập với 2 giá trị tham số thực và .

Với mỗi gia trị của và , thì hàm là một bản sao của hàm được dịch đi đơn vị trên trục thời gian. Do đó, làm tham số dịch. Nếu đặt = 0, thì ta thu được:

Như vậy, ở đây đóng vai trò là tham số tỷ lệ. Nếu thì hàm Wavelet sẽ được trải rộng, còn khi thì hàm sẽ được co lại.

Còn với phép biến đổi ngược của hàm Wavelet liên tục. Gọi

1. Biến đổi wavelet rời rạc:

Việc tính toán các hệ số Wavelet tại tất cả các tỷ lệ là một công việc hết sức phức tạp. Nếu tính toán như vậy, sẽ tạo ra một lượng dữ liệu khổng lồ. Để giảm thiểu công việc tính toán, người ta chỉ chọn ra một tập nhỏ các giá trị tỉ lệ và các vị trí để tiến hành tính toán. Hơn nữa, nếu việc tínht oán được tiến hành tại các tỷ lệ và các vị tri trên cơ sở lũy thừa cơ số 2 thì kết quả thu được sẽ hiệu quả và chính xác hơn rất nhiều. Quá trình chọn tỷ lệ và các vị trí để tính toán như trên tạo thành lưới nhị tố (dyadic), Một phân tích như trên hoàn toàn có thể thực hiện được nhờ Wavelet rời rạc

1. Biến đổi wavelet phức
2. Biến đổi wavelet nhanh

# Wavelet trên ảnh độ sâu

1. Ảnh độ sâu
2. Ảnh màu-độ sâu

# Ứng dụng của Wavelet

1. Nén ảnh JPEG 2000

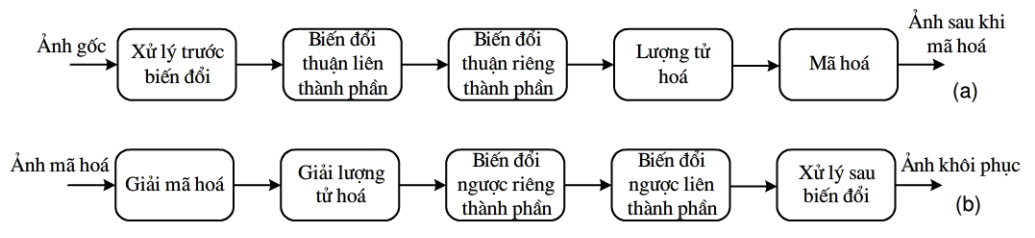
Có nhiều kỹ thuật mã hóa ảnh, ngày nay mã hóa băng con là phương pháp thành công nhất. Mã hóa băng con sử dụng wavelet tránh được hiệu ứng blocking ở tốc độ bit trung bình, bởi vì các hàm cơ sở của nó có chiều dài thay đổi. Các hàm cơ sở dài biểu diễn tín hiệu tần số thấp, các hàm cơ sở ngắn biểu diễn tín hiệu tần số cao.

Sự ra đời của JPEG đã mang lại nhiều lợi ích to lớn. JPEG có thể giảm thiểu kích thước ảnh, thời gian truyền và chi phí xử lý ảnh trong khi vẫn giữ được chất lượng ảnh khá tốt. Để việc nén ảnh có hiệu quả hơn, ủy ban JPEG đã đưa ra một chuẩn nén ảnh mới là JPEG 2000. JPEG 2000 ứng dụng biến đổi wavelet và các phương pháp mã hóa đặc biệt để có thể nén ảnh tốt hơn JPEG.

Dưới đây là bảng so sánh các tính năng của JPEG 2000 với các chuẩn nén ảnh tĩnh phổ biến:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | JPEG 2000 | JPEG-LS | JPEG | MPEG-4VTC | PNG |
| Khả năng nén ảnh không tổn thất | ★★★ | ★★★★ | ★ |  | ★★★ |
| Khả năng nén ảnh tổn thất | ★★★★★ | ★ | ★★★ | ★★★★ |  |
| Khả năng lũy tiến trong khôi phục ảnh | ★★★★★ |  | ★★ | ★★★ | ★ |
| Kỹ thuật mã hóa theo vùng | ★★★ |  |  | ★ |  |
| Khả năng tương tác với các vật thể có hình dạng bất kỳ |  |  |  | ★★ |  |
| Khả năng truy cập dòng bit ngẫu nhiên của ảnh nén | ★★ |  |  |  |  |
| Tính đơn giản | ★★ | ★★★★★ | ★★★★★ | ★ | ★★★ |
| Khả năng phục hồi lỗi | ★★★ | ★★ | ★★ | ★★★ | ★ |
| Khả năng thay đổi tỷ lệ nén | ★★★ |  |  | ★ |  |
| Tính mềm dẻo | ★★★ | ★★★ | ★★ | ★★ | ★★★ |

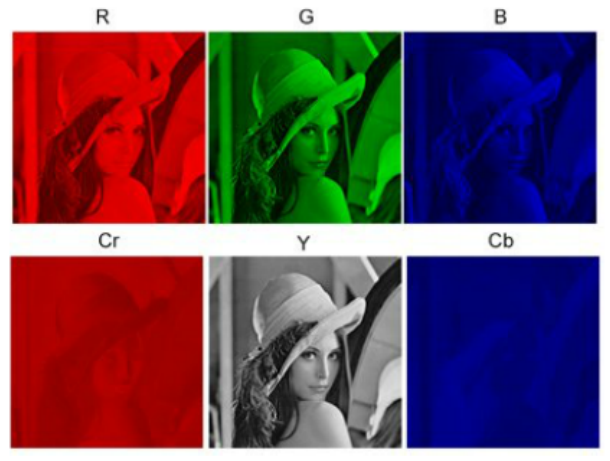
Các bước nén ảnh theo chuẩn JPEG 2000:



**Hình 1:** trình tự mã hóa (a) và giải mã JPEG 2000 (b)

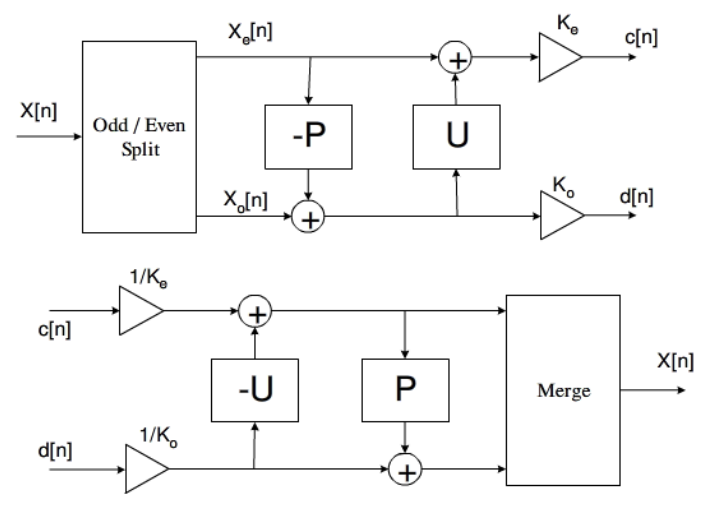
Xử lý trước biến đổi: do sử dụng wavelet, JPEG 2000 cần sử dụng ảnh đầu vào ở đang đối xứng qua 0 nên cần xử lý trước khi biến đổi. Ở phần giải mã, giai đoạn xử lý sau biến đổi sẽ trả lại giá trị gốc ban đầu cho dữ liệu ảnh.

Biến đổi liên thành phần: loại bỏ tính tương quan giữa các thành phần của ảnh. JPEG 2000 sử dụng biến đổi màu thuận nghịch (RCT) làm việc với giá trị nguyên và biến đổi màu không thuận nghịch (ICT) làm việc với giá trị thực. RCT và ICT chuyển dữ liệu ảnh từ không gian màu RGB sang YCrCb.



**Hình 2:** minh họa ảnh RGB và YCrCb

Biến đổi riêng thành phần (biến đổi wavelet): do wavelet biến đổi băng con nên các thành phần sẽ được chia thành các băng tần số khác nhau và chúng sẽ được mã hóa riêng lẻ. JPEG 2000 sử dụng biến đổi wavelet nguyên thuận nghịch 5/3 (IWT) và biến đổi thực không thuận nghịch Daubechies 9/7. Tính toán biến đổi thực hiện theo phương pháp Lifting.



**Hình 3:** phương pháp Lifting 1D dùng tính toán biến đổi wavelet

Việc tính toán biến đổi wavelet 2D suy ra từ wavelet 1D theo các phương pháp phân giải ảnh. Có 3 phương pháp phân giải ảnh trong JPEG 2000, nhưng được sử dụng nhiều nhất là phương pháp kim tự tháp.

Do biến đổi wavelet 5/3 là biến đổi thuận nghịch nên có thể áp dụng cho nén ảnh có tổn thất và không tổn thất trong khi biến đổi 9/7 chỉ áp dụng cho nén có tổn thất.

Lượng tử hóa - giải lượng tử hóa: các hệ số sẽ được lượng tử hóa theo phép lượng tử vô hướng. Bước này giúp tỷ lệ nén cao hơn bằng cách thể hiện các giá trị biến đổi với độ chính xác tương ứng cần thiết với mức chi tiết của ảnh cần nén. Các hàm lượng tử hóa khác nhau sẽ được áp dụng cho các băng con khác nhau và được thực hiện theo biểu thức:

Trong đó là bước lượng tử, là giá trị băng con đầu vào, là giá trị sau lượng tử hóa. Trong dạng biến đổi nguyên thì . Trong dạng biến đổi thực thì sẽ được chọn tương ứng với từng băng con. Bước lượng tử cho mỗi băng phải có trong dòng bit truyền đi để phía thu có thể giải lượng tử cho ảnh. Công thể giải lượng tử là:

Trong đó r là tham số dùng để xác định dấu và làm tròn, JPEG thường chọn .

Mã hóa và kết hợp dòng dữ liệu sau khi mã hóa: JPEG 2000 có thể sử dụng nhiều phương pháp khác nhau để thu được chất lượng ảnh tốt nhất. Trong thực tế có hai phương pháp cơ sở là phương pháp SPIHT và phương pháp EZW.

Phương pháp SPIHT được thiết kế tối ưu cho truyền dẫn lũy tiến, truyền các giá trị mang thông tin quan trọng đi trước. Ngoài ra nó sử dụng kỹ thuật mã nhúng nên bộ mã hóa chỉ cần nén một lần nhưng có thể giải nén ra nhiều chất lượng khác nhau. Phương pháp EZW cho tỉ lệ nén và độ tin cậy giải mã cao. Phương pháp này áp dụng nhiều cho việc nén ảnh động, rất dễ triển khai trên máy tính do không yêu cầu phức tạp về mặt lập trình.

1. Nén video

Các tín hiệu video là các chuỗi ảnh 2D khoảng 30 khung hình trên giây, chiều mới là thời gian, mở rộng việc xử lý từ 2D thành 3D. Các chuỗi biến đổi được lượng tử hóa và mã hóa entropy, sau đó sử dụng giải thuật định lượng bit dựa trên lý thuyết méo nhịp để tìm ra sự phân bố tối ưu.

Một phương pháp khác để tiếp cận với nén video là dựa trên dự đoán chuyển động. Một khung hình thường có mối liên hệ với một khung trước và sau đó. Giả thiết có thể dự đoán được các vectơ chuyển động cho tất cả các điểm ảnh để chỉ ra nơi mà mỗi phần của bức ảnh di chuyển trong các khung tiếp theo. Khi đó đủ điều kiện để gửi khung đầu tiên đã được nén và các vectơ chuyển động. Ở dãy lọc tổng hợp, khung đầu tiên được khôi phục, các khung tiếp theo được hình thành nhờ các vectơ chuyển động. Chất lượng ảnh được khôi phục phụ thuộc vào độ chính xác của việc dự đoán các vectơ chuyển động.

Chuẩn MPEG sử dụng dự đoán ngược và xuôi để dự đoán vectơ chuyển động. Các giải thuật tương tự dựa trên biến đổi wavelet cũng đang được phát triển. Việc dự đoán chuyển động cũng rất phức tạp vì có nhiều tỷ lệ: ban đầu dự đoán chuyển động theo một tỷ lệ thô, sau đó theo các tỷ lệ tinh dần. Các vùng hỗ trợ cũng phụ thuộc vào chiều dài bộ lọc.

1. Nén âm thanh

Trong hệ thống nén thoại/audio, tín hiệu được biến đổi bằng một dãy lọc cấu trúc cây. Sự định vị tần số xấp xỉ các băng tới hạn của tai người.

Nén thoại đóng vai trò rất quan trọng để giảm thời gian truyền trong thông tin di động. Thoại bao gồm thoại có thanh và thoại không thanh. Thoại có thanh thường ở tần số thấp. Trong mã hóa CELP thì thoại có thanh được mô hình là đầu ra của một bộ lọc IIR với đầu vào là nhiễu trắng. Các hệ số lọc được tìm ra nhờ việc dự đoán tuyến tính. Bộ lọc này biểu diễn hàm truyền của vùng âm thanh. Thoại không thanh có các thành phần ở tất cả các dải tần số và tương đồng với nhiễu trắng.

Về nén audio, xét một tín hiệu âm thanh lấy mẫu ở tốc độ 44,1 kHz với độ phân giải 16 bit, tốc độ bit tổng là 705,6 kb/s. Với các ứng dụng đa phương tiện thì cần phải nén trong phạm vi từ 64 đến 192 kb/s (11:1 đến 4:1). Từ việc nén audio cho thấy không có hiện tượng hao hụt trong tín hiệu được khôi phục.

1. Lưu trữ vân tay điện tử

Trong nhiều năm trước, FBI đã lưu trữ vân tay trên giấy ở một tòa nhà được bảo vệ ở Washington. Việc so khớp và truyền thông tin đi được thực hiện thủ công nên tốn rất nhiều thời gian.

Khi chúng ta xem một vân tay là một bức ảnh, ý tưởng ban đầu là tách nó thành từng khối, chẳng hạn như 256x256 điểm ảnh và gán một độ xám từ 0 (trắng hoàn toàn) đến 256 (đen hoàn toàn). Bằng cách này, chúng ta giữ lại được thông tin cần thiết bằng các chuỗi số gồm số điểm ảnh và độ xám tương ứng.

Chuỗi này có thể được lưu trữ và truyền bằng đường điện tử, và cũng có thể được so sánh với mẫu vân tay mới rất nhanh chóng. Tuy nhiên mỗi mẫu tốn khoảng 10Mb mà FBI lại có đến 30 triệu mẫu.

FBI đã nghiên cứu và sử dụng wavelet trục trực giao kép của Daubechies; ở đây là đa thức riêng phần và phù hợp với các hệ số. Trong một số trường hợp cụ thể, hóa ra chỉ cần 8% lượng thông tin gốc là đủ, phương pháp này được gọi tắt là WSQ. Nhà toán học Chris Brislawn của phòng thí nghiệm quốc gia Los Alamos New Mexico đã viết giải thuật này, đạt đến tỷ lệ nén 20:1.

1. Chứng thực vân tay

Chứng thực vân tay là một trong những phương pháp định danh cá nhân đáng tin cậy nhất và nó đóng một vai trò rất quan trọng trong các ứng dụng thường ngày và lĩnh vực pháp lý. Tuy nhiên chứng thực thủ công tốn rất nhiều thời gian và khó có thể đáp ứng yêu cầu hiện nay. Vì vậy một hệ thống chứng thực vân tay tự động (AFIS) trở thành nhu cầu tất yếu. Ở Singapore, một hệ thống an ninh mới đã được giới thiệu ở tháp Hitachi vào năm 2003: 1500 nhân viên có thể vào tòa nhà bằng cách chứng thực vân tay. Máy quét dùng tia hồng ngoại để phát hiện các hemoglobin trong máu, từ đó thu được mẫu vân trên bàn tay. Sau khi so sánh với dữ liệu đã có, hệ thống sẽ quyết định rằng người này có được quyền vào hay không.

1. Giảm nhiễu

Việc giảm nhiễu ở ảnh giúp cho việc kiểm tra và xử lý ảnh dễ dàng hơn. Lý thuyết wavlet được nghiên cứu nhiều và cho thấy nó có tác dụng hiệu quả trong việc giảm nhiễu so với nhiều phương pháp khác. Biến đổi wavelet phân chia các thành phần tham số của tín hiệu thành các dải con, được biến đổi thành nhiều mức trong khi vẫn duy trì sự định vị tín hiệu.

Trong nhiều phương pháp loại nhiễu dựa trên biến đổi wavelet, biến đổi wavelet được thực hiện, các hệ số wavelet được xử lý và các hệ số đã được xử lý sẽ được biến đổi trở lại ảnh kết quả.

Biến đổi wavelet đang dần trở thành một công cụ mạnh để loại bỏ nhiễu trong tín hiệu. Phép biến đổi phân tích tần số của tín hiệu biểu diễn trong miền tín hiệu gốc. Các phương pháp loại nhiễu được xây dựng và áp dụng cho các cơ sở lý thuyết xấp xỉ của biến đổi wavelet.

1. Các ứng dụng khác

Wavelet là công cụ rất mạnh được áp dụng trong nhiều lĩnh vực như: xử lý tín hiệu, nén dữ liệu, làm mượt và giảm nhiễu ảnh, chứng thực vân tay, phân tích protein, DNA, huyết áp và nhịp tim, mô tả lưu lượng internet, nhiều lĩnh vực vật lý như vật lý thiên văn, nhiễu động, cơ lượng tử.

Wavelet còn được áp dụng thành công ở nhiều lĩnh vực khác như địa vật lý. Chẳng hạn, wavelet trực chuẩn được áp dụng vào nghiên cứu nhiễu động ở tầng khí quyển. Theo một nghiên cứu của J.F. Howell và L. Mahrt, đo lường nhiễu động diễn ra hơn chín tiếng và sau đó được phân tích bằng phân giải wavelet. Trong một nghiên cứu của Brunet và Collineau, dữ liệu nhiễu động được ghi lại suốt một vụ mùa bắp và được phân tích bằng biến đổi wavelet.

Wavelet cũng được dùng để đo độ sâu hay địa hình của tầng nước biển. Trong một nghiên cứu của Sarah Little, việc sử dụng phân tích wavelet cho thấy mẫu và cấu trúc của dữ liệu thô.

Một ứng dụng khác thực hiện bởi Wickerhauser là ứng dụng wavelet để tìm khối u ung thư trong ảnh. Vấn đề rất lớn của ảnh y khoa là dung lượng, vì vậy nó cần được nén lại. Cũng do đó mà wavelet còn được ứng dụng vào nén dữ liệu 3-D MRI. Nén dữ liệu là ứng dụng lớn nhất của wavelet.

# Chương trình minh họa

1. Tập dữ liệu ảnh
2. Môi trường lập trình
3. Mô tả ứng dụng

# Tổng kết

# Tài liệu tham khảo

Tiếng Anh:

1. R. C. Gonzalez & R. E. Woods. *Digital Image Processing Third Edition*, pp. 483-543, 2006.
2. Richard Szeliski. *Computer Vision: Algorithms and Applications*, pp. 154-160, 2010.
3. M. Sifuzzaman, M.R. Islam, M.Z. Ali. *Application of Wavelet Transform and its Advantages Compared to Fourier Transform*, 2009.

Tiếng Việt:

1. Nguyễn Thị Lụa. *Nghiên cứu lý thuyết wavelet trong xử lý tín hiệu*, 2001.
2. Trần Duy Hưng. *Kỹ thuật xử lý ảnh sử dụng biến đổi Wavelet*.

Website: