**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 9:

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** |
| CWT | Continuous wavelet transform | Biến đổi wavelet liên tục |
| DoF | Depth of field | Độ sâu trường ảnh |
| DWT | Discrete wavelet transform | Biến đổi wavelet rời rạc |
| EZW | Embedded Zerotree Wavelet | Wavelet cây zero |
| FWT | Fast wavelet transform | Biến đổi wavelet nhanh |
| MRA | Multiresolution analysis | Phân tích đa phân giải |
| SPIHT | Set partitioning in hierarchical tree | Mã hóa phân cấp theo vùng |
| WTM | Wavelet Transform Modulus | Môđun biến đổi wavelet |
| WTMM | Wavelet Transform Modulus Maxima | Môđun biến đổi wavelet tối đa |

# Danh mục các hình

[**Hình 1:** Dạng sóng wavelet 4](#_Toc401859081)

[**Hình 2:** (Short time) Fourier Transform và Wavelet 5](#_Toc401859082)

[**Hình 3:** Lát cắt tần số-thời gian của biến đổi Fourier và biến đổi wavelet 7](#_Toc401859083)

[**Hình 4:** Daubechies wavelet 8](#_Toc401859084)

[**Hình 5:** Mexican hat wavelet 8](#_Toc401859085)

[**Hình 6:** Morlet wavelet 8](#_Toc401859086)

[**Hình 7:** Shannon wavelet 9](#_Toc401859087)

[**Hình 8:** Haar wavelet 9](#_Toc401859088)

[**Hình 9:** Minh họa lưới dyadic cho m, n 12](#_Toc401859089)

[**Hình 10:** Bộ lọc wavelet 13](#_Toc401859090)

[**Hình 11:** Dãy phân tích FWT 15](#_Toc401859091)

[**Hình 12:** (a) Dãy phân tích FWT hai giai đoạn (b) Đặc điểm phân tách tần số của nó 15](#_Toc401859092)

[**Hình 13:** Hệ số lọc Haar trực chuẩn cho 16](#_Toc401859093)

[**Hình 14:** Tính FWT hai tỷ lệ của chuỗi {1, ,4, -3, 0} dùng tỷ lệ Haar và vectơ wavelet 17](#_Toc401859094)

[**Hình 15:** Dãy lọc tổng hợp FWT-1 18](#_Toc401859095)

[**Hình 16:** Dãy lọc FWT-1 hai cấp 19](#_Toc401859096)

[**Hình 17:** Tính FWT đảo cho với tỷ lệ Haar và hàm wavelet 19](#_Toc401859097)

[**Hình 18:** Các ô (tile) tần số-thời gian của các hàm cơ sở với (a) dữ liệu mẫu, (b) FFT và (c) FWT. 20](#_Toc401859098)

[**Hình 19:** Giàn bộ lọc phân tích của FWT 2-D 22](#_Toc401859099)

[**Hình 20:** Kết quả phân giải FWT-2D 22](#_Toc401859100)

[**Hình 21:** Giàn bộ lọc tổng hợp của FWT 2-D 23](#_Toc401859101)

[**Hình 22:** Cây phân giải gói wavelet cấp 3 23](#_Toc401859102)

[**Hình 23:** Gói wavelet Haar 24](#_Toc401859103)

[**Hình 24:** Các gói wavelet được tổ chức thành cây 25](#_Toc401859104)

[**Hình 25:** Trái trên: ảnh gốc, phải trên: môđun biến đổi wavelet, trái dưới: pha môđun biến đổi wavelet, phải dưới: môđun biến đổi wavelet tối đa với vectơ pha. 27](#_Toc401859105)

[**Hình 26:** Sơ đồ khối của giải thuật đề xuất 28](#_Toc401859106)

[**Hình 27:** (a) và (b) Cặp ảnh Bản đồ gốc, (c) chênh lệch ground truth, (d) chênh lệch đã tính chưa lấp, (e) chênh lệch đã tính được lấp, (f) lỗi chênh lệch, (g) phần lấp tính từ ground truth, (h) phần lấp tính được. 30](#_Toc401859107)

[**Hình 28:** (a) Mô hình độ sâu 3D tham chiếu, (b) biên bề mặt. 31](#_Toc401859108)

[**Hình 29:** (a) và (b) Cặp stereo của các ảnh gốc, (c) bản đồ chênh lệch 2D ước lượng, (d) độ sâu 3D của mô hình. 32](#_Toc401859109)

[**Hình 30:** Biểu đồ lỗi giữa độ sâu 3D thực và cái đã phục hồi. 33](#_Toc401859110)

[**Hình 31:** Trình tự mã hóa (a) và giải mã JPEG 2000 (b) 35](#_Toc401859111)

[**Hình 32:** Minh họa ảnh RGB và YCrCb 35](#_Toc401859112)

[**Hình 33:** Phương pháp Lifting 1D dùng tính toán biến đổi wavelet 35](#_Toc401859113)

Mục lục

[Thông tin nhóm 1](#_Toc401859114)

[Danh mục các kí hiệu, chữ viết tắt và ý nghĩa 1](#_Toc401859115)

[Danh mục các hình 2](#_Toc401859116)

[I. Giới thiệu 4](#_Toc401859117)

[1) Tổng quan về wavelet 4](#_Toc401859118)

[2) Một số loại wavelet 7](#_Toc401859119)

[II. Wavelet trên ảnh màu RGB 10](#_Toc401859120)

[1) Phân tích đa phân giải 10](#_Toc401859121)

[2) Biến đổi wavelet liên tục 10](#_Toc401859122)

[3) Biến đổi wavelet rời rạc 11](#_Toc401859123)

[4) Biến đổi wavelet nhanh 13](#_Toc401859124)

[5) Biến đổi wavelet 2-D 20](#_Toc401859125)

[6) Gói wavelet 23](#_Toc401859126)

[III. Wavelet trên ảnh độ sâu 25](#_Toc401859127)

[1) Giới thiệu 25](#_Toc401859128)

[2) Môđun biến đổi wavelet 26](#_Toc401859129)

[3) Ước lượng chênh lệch 27](#_Toc401859130)

[4) Thực nghiệm 30](#_Toc401859131)

[5) Kết luận 33](#_Toc401859132)

[IV. Ứng dụng của Wavelet 34](#_Toc401859133)

[1) Nén ảnh JPEG 2000 34](#_Toc401859134)

[2) Nén video 36](#_Toc401859135)

[3) Nén âm thanh 37](#_Toc401859136)

[4) Lưu trữ vân tay điện tử 37](#_Toc401859137)

[5) Chứng thực vân tay 37](#_Toc401859138)

[6) Giảm nhiễu 38](#_Toc401859139)

[7) Các ứng dụng khác 38](#_Toc401859140)

[V. Chương trình minh họa 38](#_Toc401859141)

[1) Tập dữ liệu ảnh 38](#_Toc401859142)

[2) Môi trường lập trình 39](#_Toc401859143)

[3) Mô tả ứng dụng 39](#_Toc401859144)

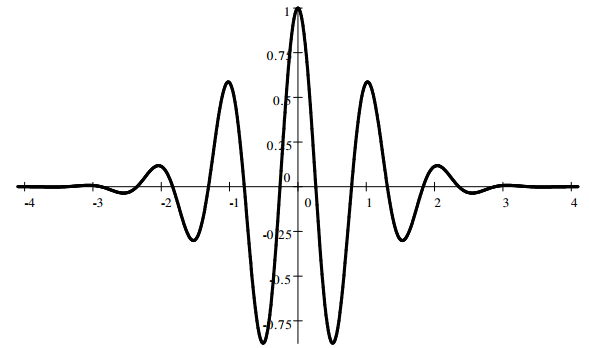
[VI. Tổng kết 45](#_Toc401859145)

[VII. Tài liệu tham khảo 45](#_Toc401859146)

# Giới thiệu

1. Tổng quan về wavelet

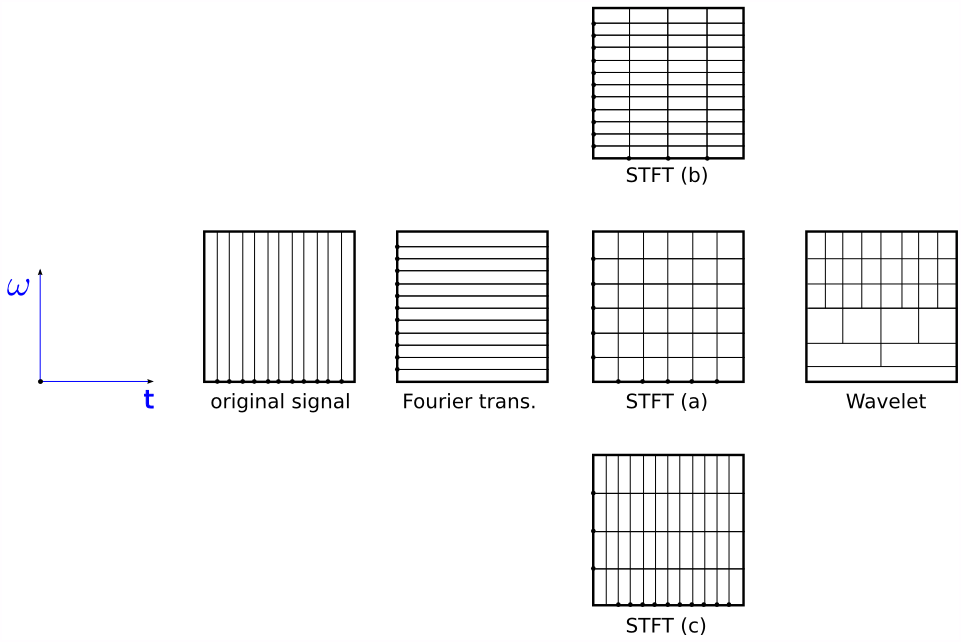
Trong toán học, một dãy wavelet là một hàm số biểu diễn bình phương khả tích (giá trị thực hoặc giá trị phức) bằng một dãy trực chuẩn tạo ra bởi một sóng wavelet. Ngày nay, biến đổi wavelet là một trong những phương pháp biến đổi mạnh nhất trong biến đổi thời gian - tần số.



**Hình 1:** Dạng sóng wavelet

Biến đổi Fourier là biến đổi được sử dụng rộng rãi trong nhiều ngành khoa học và kỹ thuật. Biến đổi Fourier chuyển một hàm tín hiệu từ miền thời gian sang tần số, nhờ đó chúng ta có thể biết được tín hiệu có các thành phần tần số nào. Tuy nhiên biến đổi Fourier có một điểm yếu là với một tín hiệu thì chúng ta không thể biết được tại một thời điểm *t*, tín hiệu đó có các thành phần tần số nào. Biến đổi wavelet ra đời đã khắc phục được điểm yếu đó. Dù chỉ làm việc với tín hiệu một chiều nhưng sau khi biến đổi xong, ta thu được một hàm số hai biến minh họa các thành phần tần số khác nhau của tín hiệu tại thời điểm *t*.

Đầu tiên, chúng ta có tín hiệu nguồn trong miền thời gian là . Với tín hiệu này, chúng ta có đầy đủ thông tin về mặt thời gian (tại thời điểm *t* thì mức năng lượng tương ứng là .



**Hình 2:** (Short time) Fourier Transform và Wavelet

Biến đổi Fourier của tín hiệu này là cho ta đầy đủ thông tin trong miền tần số nhưng lại hoàn toàn không có bất kỳ thông tin gì về miền thời gian. Với tần số thì độ lớn của nó trong tín hiệu là nhưng chúng ta không biết là tần số này xuất hiện vào thời điểm nào trong miền thời gian. Như vậy sau biến đổi Fourier, ta có thông tin trong miền tần số nhưng mất hoàn toàn thông tin về miền thời gian.

Do đó phương pháp Short-time Fourier transform (STFT) xuất hiện. Ý tưởng chính của STFT là bỏ đi một ít thông tin về các tần số thấp trong miền tần số để có thêm thông tin về miền thời gian. STFT được biểu diễn bằng một hàm theo hai biến là tần số và thời gian *t*. Như vậy nhìn vào kết quả của STFT, ta có thể biết là tần số xuất hiện vào thời điểm nào trong miền thời gian.

STFT phổ biến nhất là biến đổi Gabor. Trong biến đổi Gabor, chúng ta chọn một hàm kernel (hay còn gọi là bộ lọc), thông dụng nhất là hàm Gaussian rồi trượt hàm kernel này trong miền thời gian, thực hiện phép tích chập (convolution), rồi biến đổi Fourier. Như vậy mỗi bộ lọc được đặc trưng bằng hai tham số là vị trí và độ rộng của bộ lọc đó trong miền thời gian (tương ứng với trung bình và phương sai nếu dùng bộ lọc Gaussian).

Như vậy sẽ dẫn đến câu hỏi là chọn vị trí và độ rộng của bộ lọc như thế nào cho hợp lí. Ta chọn tham số này dựa trên trade-off sau:

* Độ rộng của bộ lọc càng lớn thì càng có thêm thông tin về tần số (thấp) nhưng mất thông tin trong miền thời gian như hình STFT(b).
* Độ rộng của bộ lọc càng nhỏ thì mất thông tin về tần số (thấp) nhưng có thêm thông tin trong miền thời gian như hình STFT(c).

Trong xử lí tín hiệu số, có đến hàng trăm loại bộ lọc khác nhau đã được đề xuất, nhưng nói chung là không có bộ lọc nào là tối ưu. Tất cả đều phải được xác định thủ công dựa vào dữ liệu thực tế.

Để khắc phục hạn chế này, người ta nghĩ ra wavelet. Ý tưởng chính của wavelet là chia miền tần số ra thành nhiều vùng và sử dụng nhiều filter với độ rộng khác nhau. Filter có độ rộng nhỏ thì chỉ bắt được các tần số cao, nhưng lại cung cấp thông tin đầy đủ trong miền thời gian. Filter có độ rộng lớn thì bắt được tần số thấp nhưng lại mất thông tin về miền thời gian. Vậy nên một cách nôm na, Wavelet transform chia miền tần số ra thành nhiều vùng. Trong vùng tần số cao (do dùng filter hẹp) ta có đầy đủ thông tin hơn về miền thời gian, ngược lại trong miền tần số thấp (filter rộng), ta mất thông tin về miền thời gian nhưng lại có thêm thông tin về các tần số thấp.

Một hàm được gọi là wavelet trực chuẩn nếu nó có thể được dùng để định nghĩa một cơ sở Hilbert (hệ trực chuẩn hoàn chỉnh) cho không gian Hilbert của các hàm bình phương khả tích.

Cơ sở Hilbert được xây dựng từ các hàm bằng trung bình tịnh tiến và giãn theo cặp của .

|  |  |
| --- | --- |
|  | (1.1-1) |

với .

Họ các hàm này là hệ trực chuẩn nếu nó là trực chuẩn trong tích nội trên .

|  |  |
| --- | --- |
|  | (1.1-2) |

trong đó là Kronecker delta.

Chúng thỏa tính hoàn chỉnh nếu mỗi hàm có thể mở rộng trong cơ sở thành

|  |  |
| --- | --- |
|  | (1.1-3) |

Một hàm *f* được biểu diễn như vậy được gọi là một dãy wavelet. Nó hàm ý rằng một wavelet trực chuẩn là tự đối ngẫu.

Biến đổi wavelet tích phân là một hàm tích phân được định nghĩa như sau:

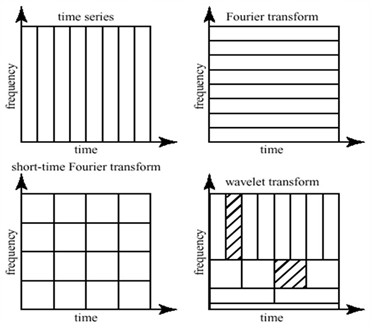
|  |  |
| --- | --- |
|  | (1.1-3) |

τ quyết định vị trí của wavelet trên miền thời gian.

s quyết định vị trí của wavelet trên miền tần số.

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

Hệ số wavelet .



**Hình 3:** Lát cắt tần số-thời gian của biến đổi Fourier và biến đổi wavelet

Công thức wavelet và biến đổi Fourier rất giống nhau. Điểm khác biệt chính là biến đổi Fourier không có tham số xác định thời gian và dùng hàm sin, cosin thay cho hàm wavelet.

STFT thời gian liên tục:

|  |  |
| --- | --- |
|  | (1.1-4) |

STFT thời gian rời rạc:

|  |  |
| --- | --- |
|  | (1.1-5) |
|  |  |

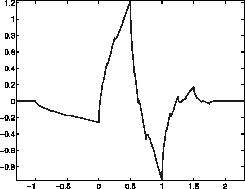
trong đó là hàm cửa sổ, là tham số thời gian, là tham số tần số, là tín hiệu cần phân tích, là nhân của biến đổi Fourier.

Trong khi đó biến đổi wavelet được biểu diễn trong miền tần số:

|  |  |
| --- | --- |
|  | (1.1-6) |

1. Một số loại wavelet

Daubechies wavelet: dựa trên nghiên cứu của Ingrid Daubechies là một họ của wavelet trực chuẩn được định nghĩa trong biến đổi wavelet rời rạc. Với mỗi loại wavelet của lớp này, có một hàm tỷ lệ (wavelet cha) tạo ra phân tích đa phân giải trực giao.

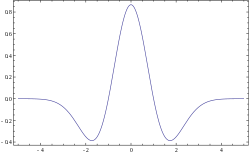


**Hình 4:** Daubechies wavelet

Biến đổi Daubechies là một trong những biến đổi phức tạp nhất trong biến đổi wavelet. Họ biến đổi này được áp dụng rất rộng rãi như nén ảnh JPEG 2000.

Mexican hat wavelet (Ricker wavelet): là một dạng đặc biệt trong họ biến đổi wavelet liên tục:

|  |  |
| --- | --- |
|  | (1.2-1) |

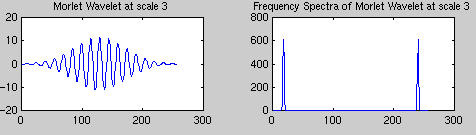


**Hình 5:** Mexican hat wavelet

Morlet wavelet:

|  |  |
| --- | --- |
|  | (1.2-2)  (1.2-3) |

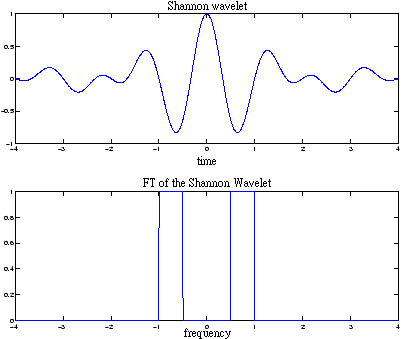
trong đó là hàm số bậc thang.



**Hình 6:** Morlet wavelet

Shannon wavelet:

|  |  |
| --- | --- |
|  | (1.2-4) |
|  | (1.2-5) |



**Hình 7:** Shannon wavelet

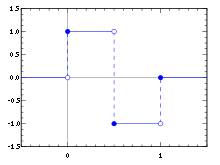
Haar wavelet:

là wavelet mẹ được định nghĩa là:

|  |  |
| --- | --- |
|  | (1.2-6) |

là hàm tỷ lệ:

|  |  |
| --- | --- |
|  | (1.2-7) |



**Hình 8:** Haar wavelet

Biến đổi Haar wavelet là biến đổi đơn giản nhất trong các phép biến đổi wavelet. Do tính chất đơn giản của Haar wavelet nên nó được ứng dụng nhiều trong nén ảnh.

# Wavelet trên ảnh màu RGB

1. Phân tích đa phân giải

Phân tích đa phân giải (MRA) là phân tích tín hiệu ở các tần số khác nhau với các độ phân giải khác nhau. Độ phân giải thời gian tốt và độ phân giải tần số kém ở tần số cao. Độ phân giải tần số tốt và độ phân giải thời gian kém ở tần số thấp. Phân tích thích hợp với tần số cao trong khoảng ngắn và thành phần tần số thấp trong khoảng dài.

Ý tưởng chính của nó là nếu một tín hiệu có thể được biểu diễn bằng tổng trọng số của thì một tập hợp lớn hơn (bao gồm cả tập ban đầu) có thể được biểu diễn bằng tổng trọng số của .

|  |  |
| --- | --- |
|  | (2.1-1) |

Không gian con .

Khi tăng kích thước không gian con sẽ làm thay đổi tỷ lệ thời gian của hàm tỷ lệ:

|  |  |
| --- | --- |
|  | (2.1-2) |

Hàm tỷ lệ là trực giao với tịnh tiến số nguyên. Các không gian con mở rộng bằng hàm tỷ lệ ở độ phân giải thấp được chưa trong các không gian con ở độ phân giải cao hơn:

|  |  |
| --- | --- |
|  | (2.1-3) |
|  | (2.1-4) |

với .

Bất kỳ hàm nào cũng được biểu diễn bằng độ chính xác tùy ý. Vì cấp của hàm mở rộng tiến đến vô cực, không gian hàm mở rộng V chứa tất cả các không gian con:

|  |  |
| --- | --- |
|  | (2.1-5) |

Một kỹ thuật xử lý ảnh nhiều tầng kha quan trọng với phân tích đa phân giải là mã hóa băng con. Trong mã hóa băng con, một ảnh sẽ được chia thành một tập gồm các thành phần giới hạn băng, gọi là băng con. Việc phân giải được thực hiện để băng con có thể tái hợp để hồi phục lại ảnh gốc mà không bị lỗi. Việc phân giải và hồi phục này được thực hiện bằng bộ lọc số.

1. Biến đổi wavelet liên tục

Biến đổi wavelet liên tục 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:

|  |  |
| --- | --- |
|  | (2.2-1) |

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

|  |  |
| --- | --- |
|  | (2.2-2) |

Ở đây, bình phương khả tích nghĩa là tích phân của bình phương của giá trị tuyệt đối 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 không. 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:

|  |  |
| --- | --- |
|  | (2.2-3) |

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 . Dấu \* ký hiệu là liên hợp phức của . Chúng ta định nghĩa hàm wavelet con theo biểu thức sau:

|  |  |
| --- | --- |
|  | (2.2-4) |

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

|  |  |
| --- | --- |
|  | (2.2-5) |

Với công thức như trên, hàm wavelet liên tục 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 .

|  |  |
| --- | --- |
|  | (2.2-6) |

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:

|  |  |
| --- | --- |
|  | (2.2-7) |

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.

Nghịch đảo CWT: khi wavelet là giá trị thực có thể tái tạo lại :

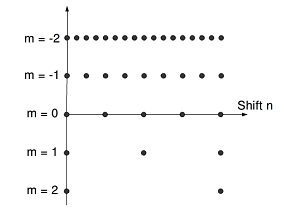
|  |  |
| --- | --- |
|  | (2.2-8) |

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 tính toán 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 biến đổi wavelet rời rạc (DWT). Do đó, việc tính toán DWT thực chất là sự rời rạc hóa biến đổi wavelet liên tục. Việc rời rạc hóa có thể thực hiện với sự lựa các hệ số τ và s như sau

|  |  |
| --- | --- |
|  | (2.3-1) |

Việc tính toán các hệ số của biến đổi Wavelet rời rạc có thể thực hiện bằng các bang lọc số nhiều nhịp đa kênh.



**Hình 9:** Minh họa lưới dyadic cho m, n

Chúng ta có tập wavelet rời rạc và tập tỷ lệ rời rạc .

Đặt

Không gian con có thể được biểu diễn bằng tổng hiệu chỉnh của các hàm mở rộng của không gian con .

|  |  |
| --- | --- |
|  | (2.3-2) |

Suy ra:

|  |  |
| --- | --- |
|  | (2.3-3) |

trong đó là hệ số hàm tỷ lệ.

Tương tự, định nghĩa

Ta có

Bất kỳ hàm wavelet nào cũng có thể được biểu diễn bằng tổng hiệu chỉnh của các hàm tỷ lệ phân giải kép dịch chuyển trong đó là hệ số hàm wavelet.

Khi áp dụng nguyên tắc mở rộng dãy, hệ số DWT của được định nghĩa là:

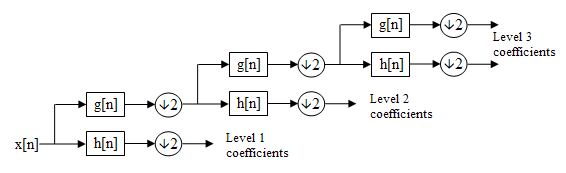
|  |  |
| --- | --- |
|  | (2.3-4) |
|  | (2.3-5) |

Trong đó là tỷ lệ tùy ý, là thừa số chuẩn hóa.

Chúng ta cũng có nghịch đảo DWT:

|  |  |
| --- | --- |
|  | (2.3-6) |

Bản chất của DWT là hệ thống các bộ lọc. Có 2 bộ lọc: bộ lọc wavelet là bộ lọc thông tần số cao và bộ lọc tỷ lệ là bộ lọc thông tần số thấp.



**Hình 10:** Bộ lọc wavelet

Trong đó g[n] là bộ lọc thông tần số thấp như hàm tỷ lệ, h[n] là bộ lọc thông tần số như hàm wavelet mẹ.

1. Biến đổi wavelet nhanh

Wavelet nhanh (fast wavelet transform – FWT) là 1 cách cài đặt hiệu quả về mặt tính toán của Wavelet rời rạc (DWT), nó khai thác mối quan hệ ngạc nhiên nhưng may mắn giữa các hệ số của DWT với các scales lân cận (adjacent scales). Còn được gọi là Mallat’s herringbone algorith, FWT giống với mô hình 2-band subband coding.

Xét công thức lọc đa phân giải:

|  |  |
| --- | --- |
|  | (2.4-1) |

Trong đó là tham số hàm vectơ và là vectơ wavelet.

Nén , dịch đi *k* và đặt , ta có:

|  |  |
| --- | --- |
|  | (2.4-2) |

Lưu ý rằng vector có thể được xem như trọng số được sử dụng để mở rộng như là tổng của các hàm nén có giá trị nén *j + 1*.

Ta lại có:

|  |  |
| --- | --- |
|  | (2.4-3) |

Biến đổi tổng và tích phân và sắp xếp lại, ta có:

|  |  |
| --- | --- |
|  | (2.4-3) |

Ta lại có:

|  |  |
| --- | --- |
|  | (2.4-4) |

Với *=j + 1* và *k = m*, ta có:

|  |  |
| --- | --- |
|  | (2.4-5) |

Lưu ý các hệ số chi tiết của nén *j* là hàm của hệ số xấp xỉ tại nén *j + 1.* Tương tự, ta có:

|  |  |
| --- | --- |
|  | (2.4-6) |

Vì các hệ số và của mở rộng dãy wavelet trở thành các hệ số và của DWT khi *f(x)* là rời rạc, ta có thể viết:

|  |  |
| --- | --- |
|  | (2.4-7) |

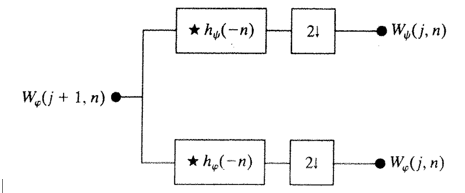
Hai công thức trên thể hiện mối liên hệ giữa các hệ số DWT của các tỷ lệ liền kề. Ta có thể thấy rằng cả và , độ nén xấp xỉ *j* và hệ số chi tiết có thể được tính bằng cách xoắn nén *j+1* hệ số xấp xỉ, với việc nén thứ tự đảo ngược và các vector wavelet và và lấy mẫu con kết quả. Vì thế chúng ta có thể viết:

|  |  |
| --- | --- |
|  | (2.4-8) |

Với độ xoắn được tính bằng đặt *n = 2k* với *k >= 0*, đánh giá độ xoắn với các chỉ số lẻ, không âm tương đương với lọc và giảm mẫu bởi 2.

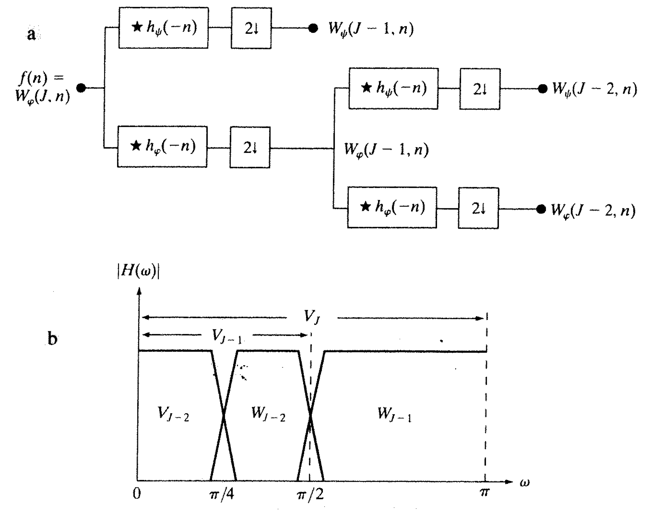
Hai công thức trên định nghĩa công thức tính cho FWT. Cho 1 chuỗi chiều dài *M = 2J*, số phép toán liên quan đến thứ tự của O(M). Nghĩa là số phép nhân và cộng tỉ lệ thuận với chiều dài cũa chuỗi nhập – vì số phép nhân và cộng liên quan đến việc quấn biểu diễn bởi dãy phân tích FWT tỉ lệ với chiều này của các chuỗi được quấn. Vì FWT so sánh với thuật toán FFT nên yêu cầu theo thứ tự các phép toán O(Mlog2M).

Để kết luận sự phát triển của FWT, chúng tôi ghi chú đơn giản dãy lọc trong hình dưới có thể được lặp để tạo ra cấu trúc nhiều giai đoạn để tính các hệ số DWT tại hai hay nhiều mức nén thành công.



**Hình 11:** Dãy phân tích FWT

Ví dụ, hình 11(a) thể hiện dãy lọc 2 giai đoạn để tạo ra cho việc tạo ra các hệ số tại 2 mức nén cao nhất của việc biến đổi. Lưu ý rằng các hệ số nén cao nhất được giả định là các mẫu của chính hàm đó. Nghĩa là với J là mức nén cao nhất.



**Hình 12:** (a) Dãy phân tích FWT hai giai đoạn  
(b) Đặc điểm phân tách tần số của nó

Dãy lọc đầu tiên trong hình 11(a) chia hàm gốc thành tần số thấp, thành phần xấp xỉ, tương ứng với hệ số nén và tần số cao, thành phần chi tiết, tương ứng với . Điều này được thể hiện trực quan trong hình 11(b), với không gian nén VJ được chia thành không gian con WJ-1 và không gian nén con VJ-1. Phổ của hàm gốc được chia thành 2 thành phần bán băng tần. Dãy lọc thứ hai trong hình 11(a) chia phổ và không gian con VJ-1, bán băng tần dưới thành các không gian con quarter-band (1/4 băng tần) WJ–2 và VJ-2 với các hệ số DWT tương ứng lần lượt là và .

Dãy lọc hai giai đoạn của hình 11(a) được mở rộng dễ dàng cho bất kỳ số nén nào. Ví dụ dãy lọc thứ ba, sẽ tính toán trên các hệ số, chi không gian nén VJ-2 thành 2 không gian con eight-band (1/8 băng tần) WJ-3 và VJ-3. Thông thường ta chọn 2J mẫu của *f(x)* và sử dụng P dãy lọc (như hình 10) để tạo ra FWT có mức nén P tại các mức nén J – 1, J – 2, .... J – P. Hệ số nén cao nhất (ví dụ J – 1) được tính đầu tiên, và nén thấp nhất (ví dụ J – P) tính cuối cùng.

Để minh họa cho các khái niệm trên, xét hàm rời rạc *f(x) = {1,4 -3,0}*. Như trong ví dụ này, chúng ta sẽ tính việc biến đổi dựa trên hàm wavelet nén Haar. Tuy nhiên, ở đây ta không sử dụng hàm cơ bản một cách trực tiếp, như đã thực hiện ở DWT. Thay vào đó ta sẽ sử dụng các vectơ và mức nén tương ứng:

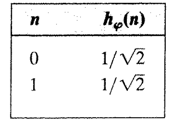
|  |  |
| --- | --- |
|  | (2.4-9) |

Đây là các hàm được sử dụng để xây dựng nên dãy lọc FWT; chúng cung cấp các hệ số lọc. Lưu ý vì nén Haar và các hàm wavelet là trực chuẩn.

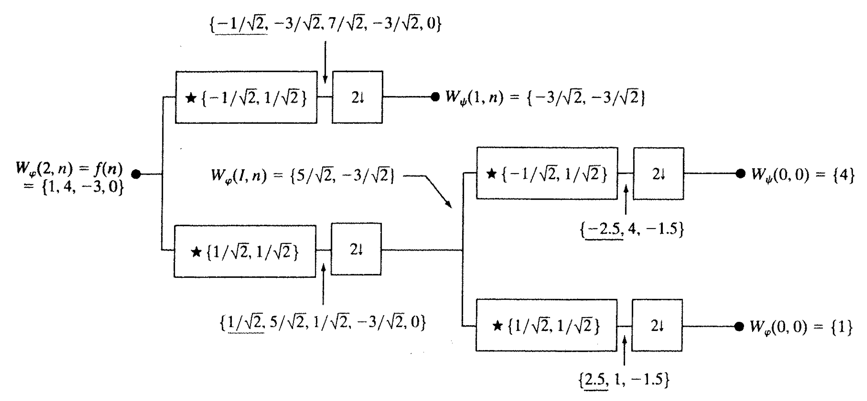
|  |  |
| --- | --- |
|  | (2.4-11) |

có thể được sử dụng để tạo ra các hệ số lọc FWT từ bộ lọc đơn nguyên mẫu (single prototype filter) – như .

Vì DWT tính từ các phần tử , ta có thể tính FWT 2 nén cho các độ nén j = {0,1}. Nghĩa là J = 2 (Có 2J =22 mẫu) và P = 2 (Ta đang làm việc với các nén J – 1 = 2 – 1 = 1 và J - P = 2 – 2 = 0 theo thứ tự đó). Việc biến đỗi sẽ được tính bằng cách sử dụng dãy lọc 2 giai đoạn của hình 11(a).



**Hình 13:** Hệ số lọc Haar trực chuẩn cho



**Hình 14:** Tính FWT hai tỷ lệ của chuỗi {1, ,4, -3, 0} dùng tỷ lệ Haar và vectơ wavelet

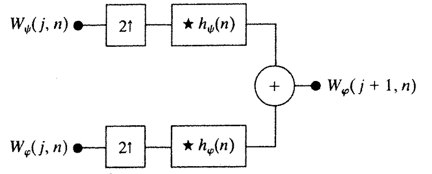
Hình 13 trình bày chuỗi các kết quả từ các giảm mẫu và FWT xoắn được yêu cầu. Lưu ý rằng hàm *f(n)* bản thân nó cũng là mức nén đầu vào (xấp xỉ) với dãy lọc phía trái cùng. Để tính hệ số xuất hiện ở cuối nhánh trên của hình 13, ví dụ đầu tiên xoắn f(n) với . Điều này đòi hỏi phải chuyển (flipping) một trong các hàm về nguyên gốc, trượt qua bên khác, và tính tổng của các sản phẩm point-wise (point-wise product) của 2 hàm. Cho chuỗi {1,4,-3,0} và , sẽ sinh ra:

Với giới hạn (term) thứ 2 tương ứng với chỉ số k = 2n = 0 (trong hình 13 các giá trị gạch dưới đại diện cho các chỉ số âm, vị dụ n < 0). Khi được giảm mẫu bằng cách lấy các điểm có chỉ số lẻ, ta lấy , cho k = {0,1}. Thay vào đó, ta có thể tính:

|  |  |
| --- | --- |
|  | (2.4-12) |

Ở đây, ta đã thay thế 2k cho n trong phép xoắn và sử dụng l như là 2 biến giả của phép xoắn (ví dụ để thay thế 2 chuỗi tương đối với nhau). Chỉ có 2 giới hạn trong tổng mở rộng vì chỉ có 2 gái trị khác 0 trong vector wavelet đảo trật tự (order-reversed wavelet vector). Thay k = 0, ta thấy ; cho k – 1, ta có Vì chuỗi đã được lọc và giảm mẫu là khớp với kết quả trước đó. Các phép xoắn và giảm mẫu còn lại được biểu diễn bằng các cách thức giống nhau.

Biến đổi ngược nhanh (fast inverse transform) để tái xây dựng *f(n)* từ các kết quả của biến đổi trước có thể được tạo thành công thức. Gọi là Inverse fast wavelet transform (FWT-1), nó sử dụng việc nén và các vectơ wavelet đã được sử dụng trong biến đổi trước, cùng với xấp xỉ mức j và các hệ số chi tiết, dể tạo ra hệ số xấp xỉ mức độ j + 1.



**Hình 15:** Dãy lọc tổng hợp FWT-1

Hình 14 mô tả chi tiết cấu trúc của nó, giống với phần tổng hợp của 2-mã hóa băng con 2-dãy và hệ thống giải mã. Như đã note ở đó, việc tái xây dựng hoàn hảo (cho các bộ lọc 2-band trực giao) yêu cầu cho i = {0,1}. Nghĩa là các bộ lọc tổng hợp và phân tích phải là phiên bản đảo ngược thứ tự của nhau. Vì các bộ lọc phân tích FWT là và bộ lọc tổng hợp FWT-1 được yêu cầu là và . Nó nên được lưu lại, tuy nhiên, điều này là cũng có thể để sử dụng phân tích nhị trực giao và các bộ lọc tổng hợp, chúng không phải là các phiên bản đảo thứ tự của nhau. Phân tích trực giao và các bộ lọc tổng hợp được biến đổi chéo qua các công thức:

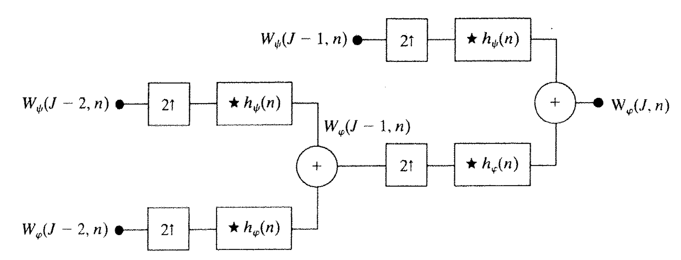
|  |  |
| --- | --- |
|  | (2.4-13) |

hoặc

|  |  |
| --- | --- |
|  | (2.4-14) |

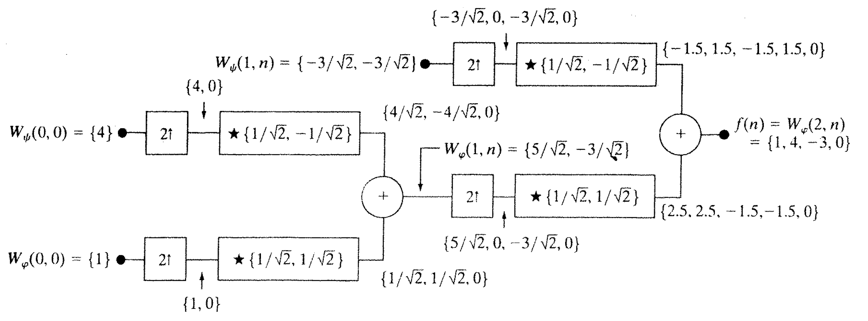
Dãy lọc FWT-1 trong hình 14 cài đặt tính toán:

với W21 đại diện cho tăng mẫu bởi 2. Các hệ số được tăng mẫu được lọc bởi phép xoắn với và và được cộng để tạo ra xấp xỉ nén lớn nhất. Thực chất, 1 xấp xỉ tốt hơn của dãy f(n) với chi tiết và phân giải tốt hơn được tạo ra. Giống với FWT trước, dãy lọc đảo có thể được lặp như trong hình 7.21, với cấu trúc nén 2 cấp để tính 2 nén cuối cùng của tái xây dựng FWT-1 được mô tả. Quá trình kết hợp hệ số có thể được mở rộng cho bất kỳ số nén nào và hứa hẹn sẽ tạo ra việc tái xây dựng hoàn hảo của chuỗi *f(n)*.



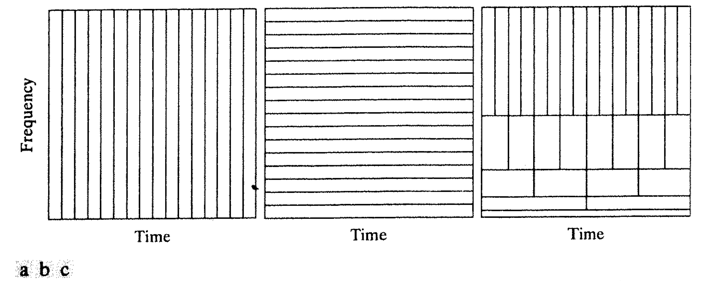
**Hình 16:** Dãy lọc FWT-1 hai cấp

Việc tính toán cho phép biến đổi wavelet nhanh đảo phản ánh bản sao trước của nó. Để bắt đầu tính toán, mức xấp xỉ 0 và các hệ số chi tiết được tăng mẫu để đạt được lần lượt là {0,1} và {4,0}. Phép xoắn với bộ lọc và sản sinh ra và , kết quả khi được cộng vào cho ra . Tiếp tục với thao tác này, f(n) được tạo ở phía phải của bank lọc tổng hợp thứ 2. Ta kết luận phần thảo luận về biến đổi wavelet nhanh bằng việc lưu ý rằng các hàm Fourier cơ bản (ví dụ sinusoids) đảm bảo sự tồn tại của FFT, sự tồn tại của FWT tuỳ thuộc vào hàm nén cho các wavelet đang được sử dụng, cũng như tính trực giao của hàm nén và các wavelet tương ứng. Vì vậy, wavelet nón Mexican (Mexican hat wavelet) vốn không có hàm nén đi kém, không thể được sử dụng trong tính toán FWT. Nói cách khác, ta không thể xây dựng bank lọc như trong hình dưới 10. cho Mexican hat wavelet, nó không thoả mãn giả thiết ngầm của cách tiếp cận FWT.



**Hình 17:** Tính FWT đảo cho với tỷ lệ Haar và hàm wavelet

Cuối cùng, ta lưu ý trong khi thời gian và tần suất thường được xem xét ở các miền khác nhau khi đại diện cho các hàm, chúng vẫn có liên kết chặt chẽ với nhau. Khi ta phân tích hàm đồng thời theo thời gian và tần suất, ta sẽ gặp vấn đề sau: nếu ta muón thông tin chính xác về thời gian, ta phải chấp nhận 1 vài sự mập mờ về tần suất. Đây là nguyên lý Heisenberg không chắc chắn (Heisenberg uncertainty principle) được áp dụng trong xử lý thông tin. Để minh hoạ cho nguyên lý này, các hàm cơ bản được sử dụng để biểu diễn cho 1 hàm có thể được xem xét dưới dạng biểu đồ như tile và mặt phẳng tần số-thời gian. Tile, còn được gọi là ô Heisenberg hay hộp Heisenberg, thể hiện nội dung tần suất của hàm cơ bản nghĩa là nó đại diện và là nơi các hàm cơ bản phụ thuộc vào thời gian.



**Hình 18:** Các ô (tile) tần số-thời gian của các hàm cơ sở với   
(a) dữ liệu mẫu, (b) FFT và (c) FWT.

Hình 17 biểu diễn các ô tần số-thời gian cho (a) hàm xung (ví dụ miền thời gian thông thường) cơ bản, (b) 1 hàm sin cơ bản (FFT), và (c) 1 FWT cơ bản. Mỗi tile là 1 vùng hình chữ nhật, chiều cao và rộng của vùng định nghĩa đặc tính về tần suất và thời gian của các hàm mà được đại diện bằng cách sử dụng hàm cơ bản. Lưu ý miền thời gian chuẩn torng hình 17(a) xác định các thời điểm khi các sự kiện xảy ra nhưng không cung cấp thông tin tần suất nào cả. Vì thế, để đại diện cho 1 tần suất hình hình đơn như sự mở rộng bằng các sử dụng hàm xung, mỗi hàm cơ bản là bắt buộc phải có. Hàm sin cơ bản trong hình 17(b), ngược lại, nó xác định các tần suất đại diện cho các sự kiện diễn ra trên 1 chu kỳ dài nhưng không cung cấp thời gian phân giải. Vì vậy, tần suất đơn hình sin được biểu diễn bởi số vô cực của các hàm xung căn bản có thể đươc biểu diễn vởi 1 sự mở rộng liên quan đến hàm sin căn bản. Tần suất và thời gian phân giải của tile FWT trong hình 17(c) lại khác, nhưng vùng của mỗi tile (hình chữ nhật) là giống nhau. Tại các tần suất thấp, tile ngắn hơn nhưng rộng hơn. Tại các nơi tần suất cao, chiều rộng của tile nhỏ hơn và cao hơn. Vì vậy các hàm FWT căn bản cung cấp sự thoả hiệp giữa 2 trường hợp giới hạn trong hình 17(a) và (b). Sự khác nhau căn bản giữa FFT và FWT đã được thảo luận trong phần giới thiệu của chương và nó rất quan trọng trong việc phân tích các hàm khôn ổn định nơi tần suất thay đổi theo thời gian.

1. Biến đổi wavelet 2-D

Biến đổi một chiều dễ dàng mở rộng thành các hàm biến đổi hai chiều như ảnh. Trong hai chiều cần dùng hàm tỷ lệ hai chiều và ba hàm wavelet hai chiều , , . Mỗi hàm là tích của hai hàm wavelet một chiều. Loại trừ tích chỉ tạo ra kết quả một chiều như , bốn tích còn lại tạo ra hàm tỷ lệ riêng lẻ

|  |  |
| --- | --- |
|  | (2.5-1) |

và các hàm wavelet riêng lẻ theo hướng

|  |  |
| --- | --- |
|  | (2.5-2) |
|  | (2.5-3) |
|  | (2.5-4) |

đo sự thay đổi theo cột (cạnh đứng), đo sự thay đổi theo dòng (cạnh ngang), đo sự thay đổi theo đường chéo.

Chúng ta có các hàm tịnh tiến và tỷ lệ cơ sở:

|  |  |
| --- | --- |
|  | (2.5-5) |
|  | (2.5-6) |

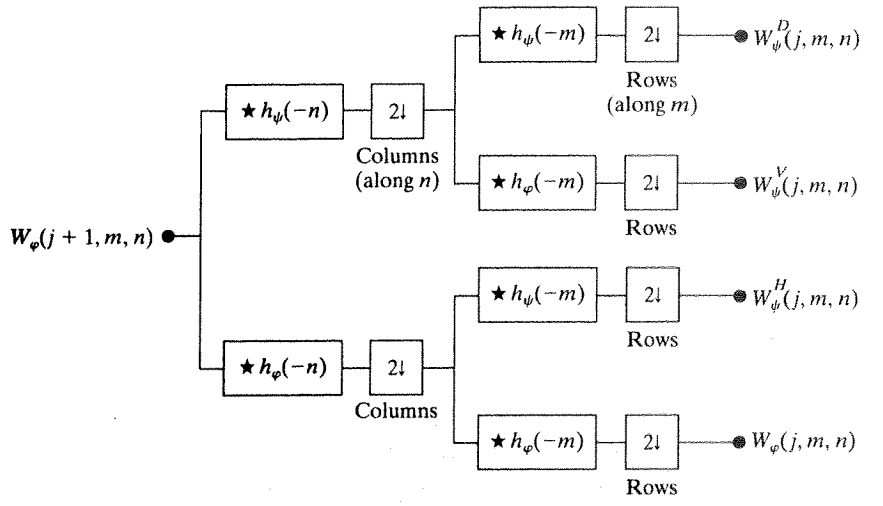
Biến đổi wavelet rời rạc của ảnh với kích thước là:

|  |  |
| --- | --- |
|  | (2.5-7) |
|  | (2.5-8) |

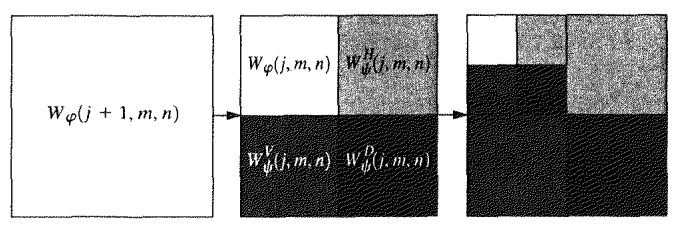
Như trong trường hợp một chiều, là tỷ lệ tùy ý và các hệ số định nghĩa một xấp xỉ ở tỷ lệ . Các hệ số bổ sung thông tin ngang, dọc, chéo cho tỷ lệ . Thông thường, chúng ta đặt và chọn với và . Với và như trên, chúng ta có nghịch đảo biến đổi wavelet rời rạc như sau:

|  |  |
| --- | --- |
|  | (2.5-9) |

Như biến đổi wavelet rời rạc một chiều, DWT 2-D có thể cài đặt bằng bộ lọc số (digital filter) và bộ giảm mẫu (downsampler). Với các hàm wavelet và tỷ lệ hai chiều tách biệt, chúng ta chỉ cần lấy 1-D FWT các dòng của , sau đó là 1-D FWT của cột kết quả.

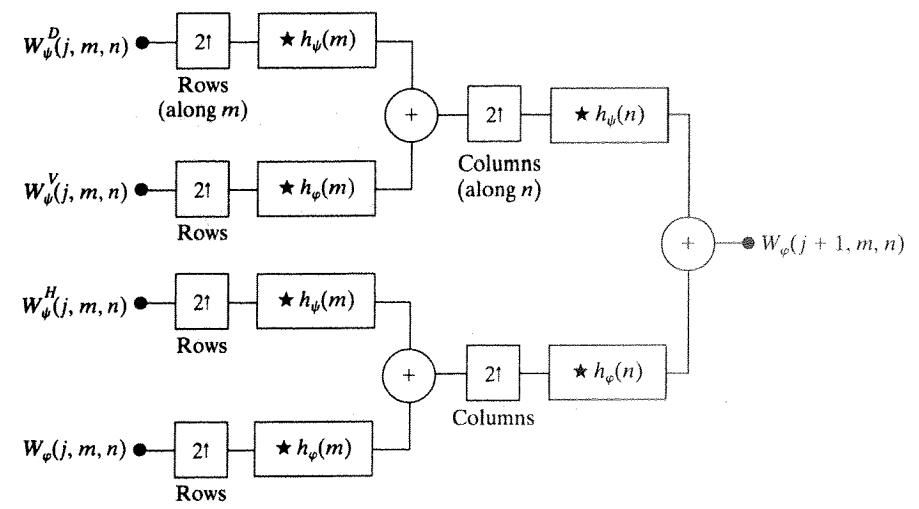


**Hình 19:** Giàn bộ lọc phân tích của FWT 2-D



**Hình 20:** Kết quả phân giải FWT-2D

Thuật toán tổng hợp tương tự như trường hợp một chiều. Ở mỗi vòng lặp, tỷ lệ xấp xỉ j và chi tiết ảnh con được tăng mẫu (upsample) và chập lại với hai bộ lọc một chiều - một bộ lọc trên cột của ảnh con và một bộ lọc trên dòng.



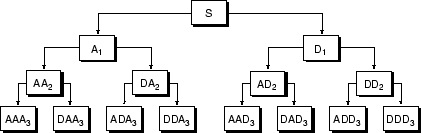
**Hình 21:** Giàn bộ lọc tổng hợp của FWT 2-D

1. Gói wavelet

Phương pháp gói wavelet là tổng quát hóa của phân giải FWT giúp phân tích tín hiệu tốt hơn. Chi phí của việc phân giải này làm độ phức tạp tính toán tăng từ của FWT lên của gói wavelet.

Trong thủ tục phân giải wavelet trực giao, bước tổng quát là tách các hệ số xấp xỉ thành hai phần. Sau đó chúng ta có được vectơ hệ số xấp xỉ và vectơ hệ số chi tiết ở tỷ lệ thô hơn. Thông tin mất giữa hai xấp xỉ liên tiếp được giữ lại trong hệ số chi tiết. Bước kế tiếp là tách vectơ hệ số xấp xỉ mới.

Trong gói wavelet, mỗi vectơ hệ số chi tiết cũng được phân giải thành hai phần với cách thức như tách vectơ xấp xỉ. Điều này giúp phân tích tốt hơn: một cây nhị phân hoàn chỉnh được tạo ra.



**Hình 22:** Cây phân giải gói wavelet cấp 3

Ý tưởng của việc phân giải là bắt đầu từ phân giải hướng tỷ lệ, sau đó phân tích tín hiệu nhận được trên băng con tần số.

Việc tính toán để tạo ra gói wavelet khá dễ dàng khi dùng wavelet trực giao. Chúng ta bắt đầu với bộ lọc kích thước , trong đó và là wavelet tương ứng.

Chúng ta định nghĩa một chuỗi các hàm

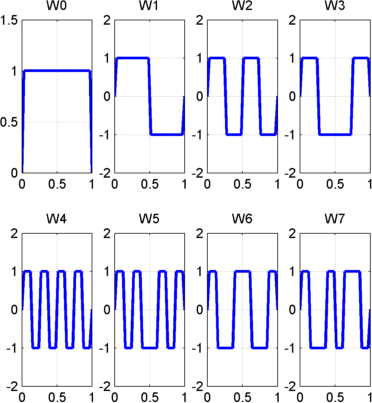
|  |  |
| --- | --- |
|  | (2.6-1) |
|  | (2.6-2) |

trong đó là hàm tỷ lệ và là hàm wavelet.

Ví dụ cho wavelet Haar, chúng ta có

Phương trình trở thành

|  |  |
| --- | --- |
|  | (2.6-3) |
|  | (2.6-4) |



**Hình 23:** Gói wavelet Haar

Tập hợp các hàm là gói wavelet . Do *j* và *n* có giá trị dương nên các gói wavelet được sắp thành cây. Tỷ lệ *j* định nghĩa độ sâu và tần số *n* định nghĩa vị trí trong cây. Chúng ta có và .



**Hình 24:** Các gói wavelet được tổ chức thành cây

# Wavelet trên ảnh độ sâu

1. Giới thiệu

Chương này sẽ trình bày về môđun biến đổi wavelet để tính toán mô hình 3D chính xác của các đối tượng. Đây là giải thuật về ước lượng độ sâu 3D cho visual inspection dùng môđun biến đổi wavelet tối đa (WTMM). Giải thuật này do Asim Bhatti, Saeid Nahavandi và Yakov Frayman [1] đề xuất.

Quá trình tái tạo 3D có thể chia làm ba mảng chính là:

* Xác định đường kính: tính toán các tham số trong và ngoài của camera.
* Tìm điểm tương ứng: tìm các cặp tương ứng của các điểm chiếu từ cùng một điểm 3D trên hai phép chiếu phối cảnh (perspective view).
* Phép đo tam giác: chiếu thông tin 2D trở lại không gian 3D để tạo ra mô hình độ sâu 3D.

Một số giải thuật được đề xuất để giải quyết các vấn đề trên trong thị giác stereo. Phần lớn các thuật toán này có thể thành hai nhóm là dựa trên bề mặt (area) hoặc dựa trên đặc trưng (feature). Hướng tiếp cận dựa trên bề mặt dựa vào mối tương quan của hai hàm ảnh trên vùng định nghĩa cục bộ. Bản đồ độ sâu dày đặc (dense depth map) có thể tìm được bằng cách tìm mối tương quan cường độ xám của vùng ảnh trong các phép chiếu, giả sử rằng vùng ảnh tương tự nhau trong các phép chiếu phối cảnh. Phương pháp dựa trên bề mặt thực hiện tốt khi có bề mặt kết cấu (texture) tốt nhưng có thể rất kém trong vùng phi đặc trưng. Một vấn đề khác ở giải thuật dựa trên bề mặt là hiệu năng kém khi xử lý biên của bề mặt, một điểm rất quan trọng trong ngữ cảnh cần tạo ra bản đồ độ sâu dày đặc và chính xác. Nói một cách khác, giải thuật dựa trên đặc trưng có gắng tạo ra sự liên kết giữa các đặc trưng được chọn, mà chúng được rút trích sử dụng các giải thuật rút trích đặc trưng rõ ràng. Hạn chế chính của nó là thông tin độ sâu nhận được rất thưa nên khó phục hồi được bản đồ độ sâu dày đặc. Nó còn có khuynh hướng bị kẹt ở các đặc trưng ảnh không cần thiết và phụ thuộc vào tính hiệu quả của giải thuật rút trích đặc trưng.

Mối quan tâm chính ở đây là sự phát triển của một giải thuật stereo vision có khả năng ước lượng chính xác độ sâu 3D của của bộ phận máy tạo khuôn rập kim loại. Mục tiêu là phát triển một hệ thống visual inspection ít dùng đến phần cứng không giống như các kỹ thuật đã có vốn dùng đến máy quét laser, laser và máy chiếu mẫu sáng. Động lực chính sau giải thuật là giữ cho hệ thống đơn giản, liên quan đến tối thiểu phần cứng và độ phức tạp để làm cho hệ thống ít tốn chi phí và không cồng kềnh. Do phụ thuộc phần cứng tối thiểu nên nó có thể cài đặt trong môi trường nhà máy khác nhau mà không cần thay đổi nhiều trong việc thiết lập.

Khi xem xét các bộ phận máy tạo khuôn rập kim loại với bề mặt sáng có kết cấu nghèo nàn và các vết không cần thiết, việc dùng giải thuật dựa trên bề mặt hay đặc trưng đều không đem lại kết quả tốt. Giải thuật dựa trên bề mặt khó ước lượng độ sâu chính xác khi thiếu kết cấu bề mặt tốt, trong khi đó giải thuật dựa trên đặc trưng lại gặp khó khăn do có các vết không cần thiết trên bề mặt. Điều này đưa chúng ta đến chiến thuật so khớp thô sang tinh (coarse to fine matching) dựa trên phân tích đa phân giải, khi đó mối liên quan được xác định ở các tỷ lệ và độ phân giải khác nhau.

Trong ngữ cảnh so khớp stereo dùng phân tích đa phân giải, một số nghiên cứu dùng wavelet vô hướng và có kết quả khá triển vọng. Trong phần này, hướng tiếp cận dựa trên multiwavelet được dùng để xác định vấn đề của so khớp ảnh stereo. Bằng việc sử dụng khái niệm của môđun biến đổi wavelet tối đa (wavelet transform modulus maxima - WTMM), hiệu năng của một số giải thuật đã được cải thiện, chính là nhờ vào việc dùng hệ số vectơ của môđun biến đổi wavelet với pha và độ lớn chỉ đến pháp tuyến của các cạnh.

1. Môđun biến đổi wavelet

Lý thuyết wavelet có thể được định nghĩa bằng hai hàm giãn nở như sau:

|  |  |
| --- | --- |
|  | (3.2-1) |
|  | (3.2-2) |

Trong đó và là các hệ số của hàm tỷ lệ và wavelet , *M* và *k* đại diện cho dãy lọc và integer shift. Hàm đa tỷ lệ và multiwavelet có thể biểu diễn thành

|  |  |
| --- | --- |
| , | (3.2-3) |

Trong đó *r* đại diện cho vô số multiwavelet đồng nhất trong trường hợp của wavelet vô hướng. Tương tự trong ngữ cảnh của multimwavelet và trong (3.2-1) và (3.2-2) là ma trận số thực của các hệ số đa bộ lọc định nghĩa hàm wavelet và tỷ lệ. Môđun biến đổi wavelet (wavelet transform modulus - WTM) có thể được biểu diễn là

|  |  |
| --- | --- |
|  | (3.2-4) |

Trong đó đại diện cho độ lớn của WTM của hệ số thứ *k*, trái lại *s* đại diện cho tỷ lệ của không gian con. Ngoài ra độ lớn của WTM có thể biểu diễn là

|  |  |
| --- | --- |
|  | (3.2-5) |

Trong đó và là không gian chi tiết ngang và dọc. Tương tự pha của WTM có thể biểu diễn là

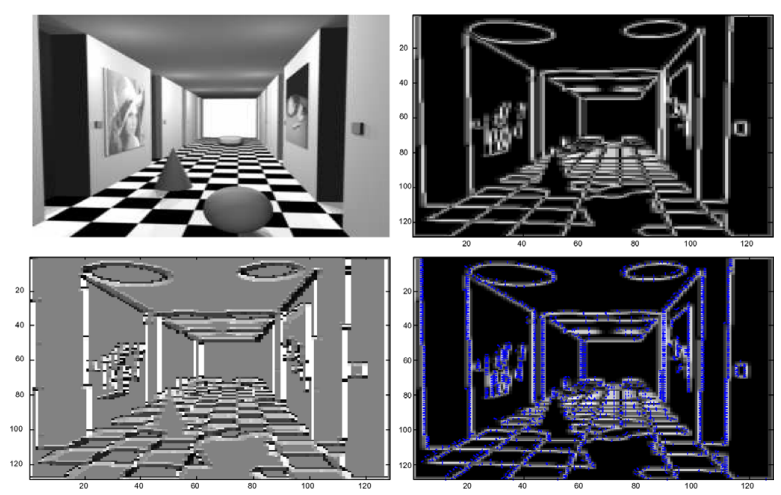
|  |  |
| --- | --- |
|  | (3.2-6) |

Trong đó

|  |  |
| --- | --- |
|  | (3.2-7) |

Vectơ chỉ đến phương pháp tuyến của bề mặt cạnh như sau

|  |  |
| --- | --- |
|  | (3.2-8) |



**Hình 25:** Trái trên: ảnh gốc, phải trên: môđun biến đổi wavelet, trái dưới: pha môđun biến đổi wavelet, phải dưới: môđun biến đổi wavelet tối đa với vectơ pha.

Một điểm cạnh là điểm *p* ở tỷ lệ *s* như tối đa cục bộ ở và với đủ nhỏ. Tối đa đó gọi là môđun biến đổi wavelet tối đa (WTMM) và cố định trượt (shift invariant). Ví dụ về môđun biến đổi wavelet trình bày ở hình trên.

1. Ước lượng chênh lệch

Quá trình so khớp bắt đầu với chuẩn hóa tỷ lệ, làm tối thiểu hiệu ứng của biến động độ sáng trong mỗi tỷ lệ vốn có thể tồn tại giữa các tỷ lệ của phép chiếu phối cảnh và có thể biểu diễn là

|  |  |
| --- | --- |
|  | (3.3-1) |

Trong đó là hệ số tỷ lệ wavelet chuẩn hóa, *i* tham chiếu đến tất cả không gian chi tiết và đại diện cho không gian xấp xỉ trong tỷ lệ *s*. Sau khi biến đổi wavelet và tính toán WTM, chúng ta sẽ có ma trận trọng số làm nổi bật được đặc trưng. Sau quá trình chuẩn hóa tỷ lệ là đo độ tương đồng dựa trên tương quan, biểu diễn là

|  |  |
| --- | --- |
| với | (3.3-2) |

Trong đó và biểu diên WTMM đã chuẩn hóa liên quan đến ảnh thứ nhất và thứ hai. Sơ đồ khối ở hình bên dưới sẽ giải thích rõ hơn về thuật toán.

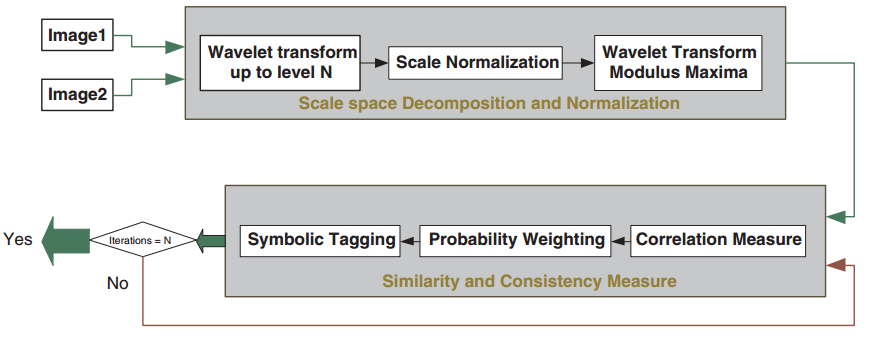
* 1. Đo xác suất

Để cho các ứng viên so khớp nhất quán, một thủ tục gán nhãn ký hiệu sẽ thực hiện dựa trên xác suất xuất hiện và ba ngưỡng . Giá trị của các ngưỡng thường trong [0.5 0.7], [0.70 0.85] và [0.85 0.95]. Giá trị của được chọn nhờ phương pháp thử sai. Xác suất xuất hiện là xác suất chọn điểm bất kỳ làm ứng viên trong không gian tìm kiếm từ các không gian được định nghĩa là

|  |  |
| --- | --- |
| trong đó | (3.3-3) |

Ở đây, là số lần một ứng viên phù hợp được chọn ra từ không gian . Tất cả ứng viên phù hợp có xác suất được chọn như nhau (). Hệ số tương quan cho mỗi ứng viên là xác suất xuất hiện, biểu diễn là

|  |  |
| --- | --- |
| trong đó | (3.3-4) |



**Hình 26:** Sơ đồ khối của giải thuật đề xuất

* 1. Gán nhãn ký hiệu

Các ứng viên sau đó sẽ chia thàn những tổ hợp dựa trên ngưỡng . Ngưỡng đầu tiên được áp dụg ngay sau bước tương quan để lọc bỏ ứng việc không tiềm năng. Các ứng viên còn lại sẽ được gán nhãn ký hiệu tùy vào tính nhất quán của nó như sau

|  |  |
| --- | --- |
|  | (3.3-5) |

Như biểu thức đầu, không có nhập nhằng ở các ứng viên với nhãn Opc vì xác suất nhất quán là 1, ngược lại nhập nhằng có thể tồn tại ở các ứng viên có nhãn Cd và Cr. Nhập nhằng là hiện tượng có hơn một ứng viên cho một điểm trong ảnh tham chiếu. Để giải quyết vấn đề nhập nhằng cần thực hiện tối ưu hình học. Các ứng viên thỏa yếu tố topo hình học và có nhãn Cd sẽ được nâng cấp thành Op, nếu có nhãn Cr sẽ nâng thành Cd. Sau đó tất cả các cặp có nhãn Op sẽ được xem là ứng viên tiềm năng, còn các ứng viên có nhãn Cd sẽ được xem xét ở cấp độ nội suy. Mục đích của việc giữ các điểm có nhãn Cd là để khai thác tiềm năng của chúng trong cấp phân giải kế tiếp. Các ứng viên không thỏa các điều kiện trên sẽ bị vứt bỏ.

* 1. Lọc hình học

Vì có không gian tìm kiếm nên có khả năng việc chọn các điểm khác từ không gian tìm kiếm sẽ bị nhập nhằng. Để giải quyết vấn đề này cần thực hiện lọc topo hình học để rút trích các ứng viên tiềm năng ra khỏi tổ hợp các ứng viên nhập nhằng. Để thực hiện, định hướng hình học của các điểm nhập nhằng tham chiếu đến Op của (3.3-5) sẽ được kiểm tra và các cặp có topo hình học gần nhất tương ứng với Op sẽ được chọn làm ứng viên tiềm năng. Ba đặc trưng hình học là sai biệt khoảng cách tương đối, sai biệt khoảng cách tuyệt đối và sai biệt độ dốc của đừng liên kết các điểm đã được chọn và tính toán để kiểm tra độ tương tự hướng hình học. Đo lường hình học sau đó sẽ được tính cùng với điểm tương quan ứng viên để giữ các viên đã được xem xét trước đó như sau

|  |  |
| --- | --- |
|  | (3.3-5) |

Trong đó , là trung bình sai biệt khoảng cách tương đối và tuyệt đối, là sai biệt độ dốc trung bình. Các đặc trưng hình học này được tính nhiều lần để tối thiểu ảnh hưởng của bất kỳ ứng viên tham chiếu sai. Ba đặc trưng hình học bất biến ma biến đổi hình học như Affine, Metric, Euclidean,… vốn rất phổ biến trong ứng vụng thị giác stereo và ước lượng độ sâu 3D. Các ứng viên có topo hình học gần nhất, nghĩa là *Gc* thấp nhất sẽ được chọn làm ứng viên tối ưu.

* 1. Nội suy tỷ lệ

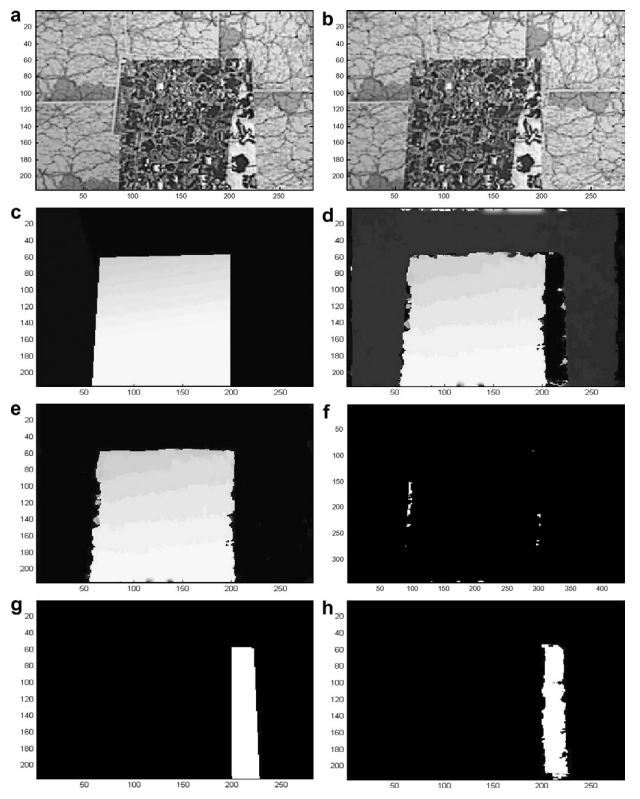
Quá trình so khớp ở cấp tho nhất kết thúc với một số các cặp cần nội suy sang cấp độ tinh hơn. Mối quan hệ chùm giữa các hệ số ở cấp thô và cấp tinh có thể được hình dung bằng cách lấy một phần mười của thừa số 2, tuy nhiên ở đây dùng trung bình ở các cặp hệ liên tục trên dòng và cột. Lí do dùng trung bình là để giữ lại ý nghĩa của tất cả hệ số đang xem xét. Ngoài ra có thể tồn tại WTMM bị mất do lấy một phần mười của thừa số 2. Độ chênh lệch từ lệch thô sang lệch tinh có thể cập nhật như sau

|  |  |
| --- | --- |
|  | (3.3-5) |

Trong đó là chênh lệch ước lượng ở mức tỷ lệ hiện tại.

1. Thực nghiệm

Thuật toán đề xuất ở đây thực chất là phát triển lên từ giải thuật multiwavelet cũng do nhóm của Asim Bhatti đề xuất, chúng ta có thể thấy rõ các cải tiến, bất chấp chất lượng của bản đồ chênh lệch (disparity map). Để đánh giá hiệu năng của giải thuật, nhóm tác giả thực hiện hai thực nghiệm khác nhau. Ở thực nghiệm thứ nhất, các giải thuật mục 3 được áp dụng cho nhiều từ trang web của trường Middlebury College. Những bức ảnh này khác nổi tiếng trong lĩnh vực thị giác máy tính do tính đa dạng về kết cấu bề mặt và cường độ.



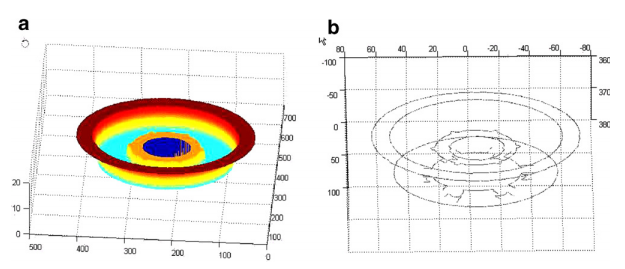
**Hình 27:** (a) và (b) Cặp ảnh Bản đồ gốc, (c) chênh lệch ground truth, (d) chênh lệch đã tính chưa lấp, (e) chênh lệch đã tính được lấp, (f) lỗi chênh lệch, (g) phần lấp tính từ ground truth, (h) phần lấp tính được.

Kiểm tra hiệu năng thống kê được tính bằng tỷ lệ phần trăm của điểm ảnh xấu B được định nghĩa là

|  |  |
| --- | --- |
|  | (3.4-1) |

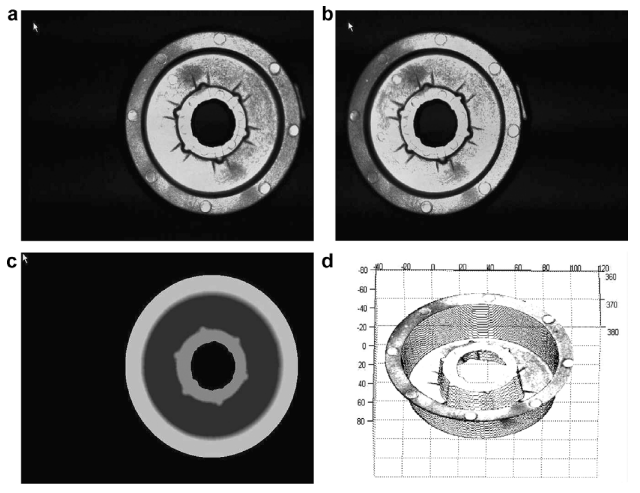
Trong đó và biểu diễn bản đồ chênh lệch ground-truth và đã tính toán, biểu diễn sai biệt chênh lệch chấp nhận được (0.5). Một nội suy tuyến tính đơn giản được thực hiện để lấp các lỗ trống trong bản đồ chênh lệch để không có giá trị chênh lệch nào, giả sử rằng chênh lệch có ở tất cả điểm không liên tục, nghĩa là ở tất cả môđun biến đổi wavelet tối đa.

Một trong nhưng kết quả thực nghiệm được trình bày như hình trên. Ngoài ra việc phát hiện trước các lỗ cần lấp không được thực hiện trong giải thuật đề xuất. Vùng cần lấp sau khi được sẽ phát hiện sẽ rút trích ra bằng việc giữ lại quá trình ghép cặp, thu được kết quả khá triển vọng mà không cần tối ưu thêm. Bất kỳ cặp nào không thỏa điều kiện trong mục 3 ở trên sẽ bị vứt bỏ. Như ở *hình 27*(f), chỉ vài cặp được tìm thấy ở vị trí lấp. Một phép nội suy tuyến tính sau đó được thực hiện. Bản đồ chênh lệch sau khi nội suy sẽ như *hình 27*(e), sau đó là lỗi pixel ở *hình 27*(f). *Hình 27*(g) và *hình 27*(h) cho thấy các hình đã lấp từ bản đồ chênh lệch trong *hình 27*(c) và hình ước lượng *27*(d). Mức lỗi của cặp pixel lỗi là 2.2%.



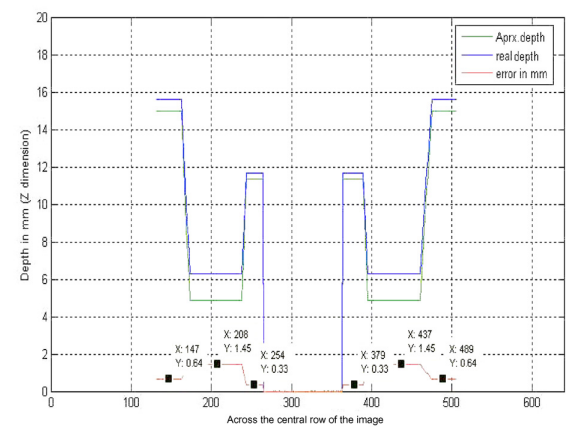
**Hình 28:** (a) Mô hình độ sâu 3D tham chiếu, (b) biên bề mặt.

Thực nghiệm thứ hai thực hiện trên bộ phận máy xe kim loại công nghiệp. Mục tiêu chính là phát triển một thuật toán có thể ước lượng chính xác độ sâu 3D của bộ phận máy tạo khuôn kim loại ít dùng đến phần cứng. Lí do chính là để sử dụng được trong nhiều môi trường công nghiệp khác nhau mà không phải thiết lập lại. Vì môi trường stereo vision chưa đủ mạnh để giải quyết bài toán ước lượng độ sâu chính xác nên cần dùng thêm thông tin từ mô hình 3D tham chiếu. Hiện tại thực nghiệm thực hiện trên vỏ ống bơm kim loại. Mô hình tham chiếu của vỏ ống bơm kim loại này được giới thiệu trong Matlab bằng cách đo thủ công các chiều. Ví dụ về nó được trình bày trong *hình 28*. Thông tin cần trong mô hình này là vị trí tương đối của các điểm không liên tục và dải chiều của độ sâu 3D, từ đó đưa ra thông tin về độ chênh lệch. Vị trí tương đối của các cạnh bề mặt khác nhau là quan trọng nhất vì giải thuật stereo vision thực hiện không tốt trên bề mặt có độ sâu đột ngột. Ngoài ra mỗi loại đối tượng đều cần một mô hình tham chiếu.



**Hình 29:** (a) và (b) Cặp stereo của các ảnh gốc, (c) bản đồ chênh lệch 2D ước lượng, (d) độ sâu 3D của mô hình.

Việc thiết lập phần cứng khá đơn giản, chỉ cần các camera stereo từ Videre Design và đèn huỳnh quang tròn. *Hình 29* trình bày một ví dụ về mô hình độ sâu ước lượng. Chúng ta có thể thấy một bản đồ chênh lệch rất mượt có được sau khi tối ưu và cải tiến bằng thông tin rút trích được từ mô hình tham khảo.



**Hình 30:** Biểu đồ lỗi giữa độ sâu 3D thực và cái đã phục hồi.

Ngoài ra mô hình này hoàn toàn độc lập với hướng hay vị trí của đối tượng trong miền ảnh. Biểu đồ lỗi được trình bày trong *hình 30*. Nó cho thấy sự khác biệt trong độ sâu của các vỏ máy bơm khác nhau, giữa mô hình tham chiếu trong *hình 28* và ước lượng trong *hình 29*. Độ lỗi nằm trong khoảng [0.65 mm 1.45 mm] là khá triển vọng trong một giải thuật đơn giản được đề xuất. Ngoài ra một số thực nghiệm khác cũng đang được tiến hành để kiểm tra tính nhất quán của giải thuật trong nhiều môi trường và điều kiện sáng khác nhau. Mục tiêu là xây dựng một hệ thống phát hiện lỗi tự động sử dụng thông tin độ sâu.

1. Kết luận

Phân tích đa phân giải dựa trên giải thuật thị giác stereo đã được trình bày trong chương này. Kỹ thuật chính là ước lượng các cặp tối ưu dẫn đến ước lượng chính xác độ sâu 3D. Giải thuật trình bày là hướng tiếp cận dựa trên phần mềm, hạn chế liên quan đến phần cứng. Hiệu quả và độ chính xác của giải thuật nằm ở mô hình tham chiếu 3D có sẵn thay vì thêm vào công cụ phần cứng như máy quét laser, laser và máy chiếu mẫu sáng. Bơỉ vì sự phụ thuộc tối thiểu của nó vào hệ thống làm tăng tính hiệu quả và có thể triển khai trong những môi trường công nghiệp khác nhau với môi trường khác nhau và các biến chiếu sáng và không có sự thay đổi đáng kể trong các tham số đề xuất của hệ thống.

Để giữ cho quá trình được thống nhất và ổn định lỗi thỉ dùng 2 kỹ thuật: đánh ký hiệu và làm mịn hình học, đã được giới thiệu và ứng dụng vào những kỹ thuật ràng buộc phổ biến và liên tục, được trình bày và áp dụng cùng với ràng buộc phổ biến về tính duy nhất và liên tục. Vấn đề nhập nhằng rõ ràng và lỗ hổng bị ẩn được trình bày sử dụng cải tiến topo hình học và gán nhãn ký hiệu. Thủ tục gán nhãn ký hiệu theo dõi ứng viên và khả năng của chúng để chia thành các tập và cấp độ khác nhau. Nhãn của các ứng viên sẽ được cập nhật với hiệu năng và tiềm năng thay vì vứt bỏ ngay.

# 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 | ★★★ | ★★★ | ★★ | ★★ | ★★★ |

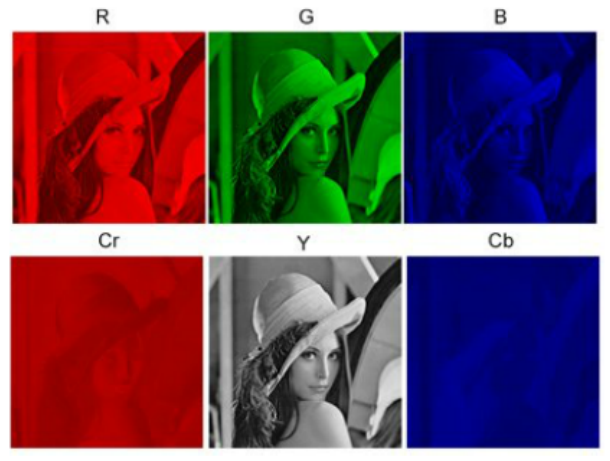
Các bước nén ảnh theo chuẩn JPEG 2000:



**Hình 31:** Trình tự mã hóa (a) và giải mã JPEG 2000 (b)

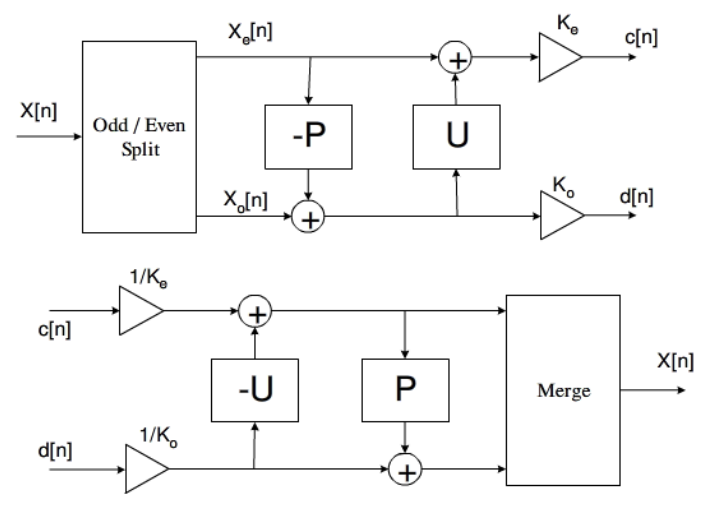
Xử lý trước biến đổi: do sử dụng wavelet, JPEG 2000 cần sử dụng ảnh đầu vào ở đang đối xứng qua 0 nên cần xử lý trước khi biến đổi. Ở phần giải mã, giai đoạn xử lý sau biến đổi sẽ trả lại giá trị gốc ban đầu cho dữ liệu ảnh.

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 32:** 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 33:** 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:

|  |  |
| --- | --- |
|  | (4.1-1) |

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:

|  |  |
| --- | --- |
|  | (4.1-2) |

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

"Standard" test images

<http://www.imageprocessingplace.com/root_files_V3/image_databases.htm>

IEEE International Conference on Robotics and Automation (ICRA)

<http://rgbd-dataset.cs.washington.edu>

VOCB3DO, Berkeley 3-D Object Dataset

<http://kinectdata.com>

NYU Depth Dataset V2

<http://cs.nyu.edu/~silberman/datasets/nyu_depth_v2.html>

1. Môi trường lập trình

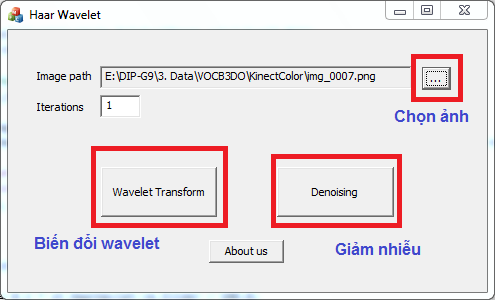
Ngôn ngữ: C++

OpenCV 2.4.8

IDE: Visual Studio 2010

1. Mô tả ứng dụng

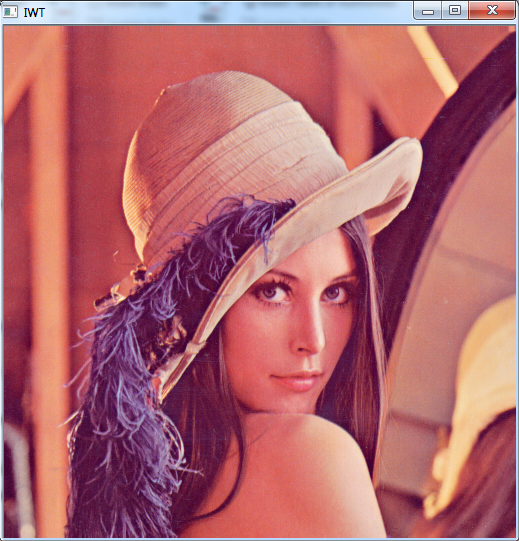
Màn hình giao diện chính



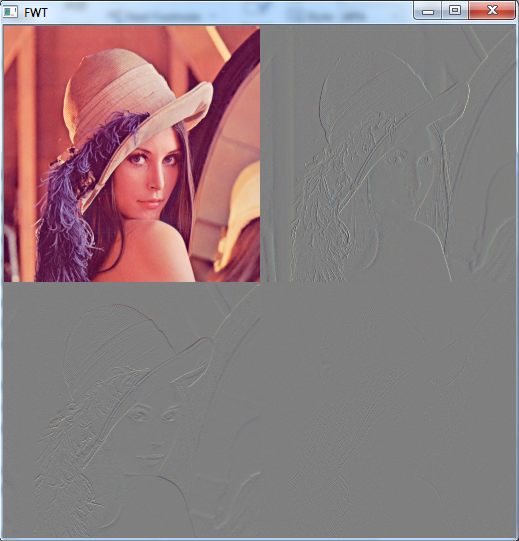
Chức năng chính

* Biến đổi wavelet với tham số iteration
* Giảm nhiễu, với mức threshold

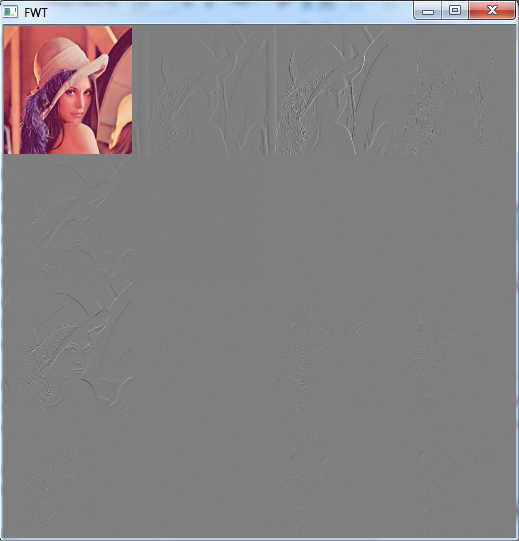
Chạy Wavelet transform với ảnh “Invert\_1\_lena\_color\_512.tif”



Ảnh sau khi biến đổi với iteration chọn ban đầu là 1.



Ảnh sau khi biến đổi với iteration chọn ban đầu là 2.



Thực hiện biến đổi wavelet trên ảnh “00000-color.png”

|  |  |
| --- | --- |
| Ảnh gốc | Wavelet transform (iteration = 3) |

Thực hiện tương tự với ảnh độ sâu “00000-depth.png” trong ICRA/ scene\_01.

|  |  |
| --- | --- |
| Ảnh gốc | Wavelet transform (iteration = 1) |
| Wavelet transform (iteration = 2) | Wavelet transform (iteration = 3) |

Thực hiện tương tự với ảnh RGB-D “00000-depth.png” trong ICRA/ scene\_01.

|  |
| --- |
| Ảnh gốc |
| Wavelet transform (iteration = 1) |
| Wavelet transform (iteration = 2) |

Chạy chức năng giảm nhiễu với ảnh “nyu\_depth\_v2\_raw.jpg”

|  |
| --- |
| Ảnh gốc, làm nhiễu tự động |
| Ảnh sau khi chỉnh threshold = 50 |

Chạy chức năng giảm nhiễu với ảnh “nyu\_depth\_v2\_raw.jpg”

|  |
| --- |
| Ảnh gốc, làm nhiễu tự động |
| Ảnh sau khi chỉnh threshold = 60 |

# Tổng kết

Như vậy, xuất phát từ nhu cầu xử lý tín hiệu thành mô tả các tín hiệu thực, để từ các tín hiệu này, có thể tính toán, nén hoặc khử nhiễu đối với dữ liệu ảnh. Có rất nhiều phương pháp để biến đổi ảnh từ miền không gian thành miền tần số đã được đề xuất và áp dụng. Trong bài báo cáo này, nhóm đã trình bày về phép biến đổi Fourier, phép biến đổi wavelet cùng với phân tích đa phân giải và mã hóa băng con. Với sự ra đời đầu tiên, phép biến đổi Fourier vẫn còn nhiều hạn chế như không thể lưu trữ được thông tin về thời gian, hoặc xử lý trên các tín hiệu động. Do đó, biến đổi wavelet đã ra đời và khắc phục được những khuyết điểm này.

Ứng dụng của biến đổi wavelet đã được trình bày ở IV, có thể thấy biến đổi wavelet có rất nhiều khả năng và ứng dụng trong thực tế. Hiện nay, hướng xử lý wavelet vẫn tiếp tục được nghiên cứu để có thể áp dụng thêm những phép tính toán, xử lý hoặc đề xuất thêm các dòng wavelet mới trong tương lai.

# Tài liệu tham khảo

Tiếng Anh:

1. Asim Bhatti, Saeid Nahavandi, Yakov Frayman: “*3D depth estimation for visual inspection using wavelet transform modulus maxima*”, Comput. Electr. Eng., Vol. 33, August 2006.
2. M. Sifuzzaman, M.R. Islam, M.Z. Ali: “*Application of Wavelet Transform and its Advantages Compared to Fourier Transform”*, Vidyasagar University, 2009.
3. Rafael C. Gonzalez & R. E. Woods: “*Digital Image Processing Third Edition”*, Prentice Hall, ISBN 978-0131687288, pp. 483-543, 2006.
4. Richard Szeliski: “*Computer Vision: Algorithms and Applications”*, Springer, ISBN 978-1848829343, pp. 154-160, 2010.

Tiếng Việt:

1. Đỗ Ngọc Anh: “Nén ảnh sử dụng biến đổi wavelet và ứng dụng trong các dịch vụ dữ liệu đa phương tiện di động”, Đại học Bách Khoa Hà Nội, 2006.
2. Nguyễn Thị Lụa: “*Nghiên cứu lý thuyết wavelet trong xử lý tín hiệu”*, Đại học Bách Khoa Hà Nội, 2001.
3. Trần Duy Hưng: “*Kỹ thuật xử lý ảnh sử dụng biến đổi Wavelet*”.

Website:

1. <http://en.wikipedia.org/wiki/Wavelet_transform>
2. <http://www.mathworks.com/help/wavelet/ug/wavelet-packets.html>