ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN
UIT

GIÁO TRÌNH

XỬ LÝ ẢNH

Biên soạn: ThS. Võ Đức Khánh
Hiệu đính: GS.TSKH. Hoàng Kiếm

NHÀ XUẤT BẢN
ĐẠI HỌC QUỐC GIA TP HỒ CHÍ MINH

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN

Giáo trình

XỬ LÝ ẢNH

Tác giả: ThS Võ Đức Khánh

Hiệu đính: GS-TSKH Hoàng Văn Kiếm

NHÀ XUẤT BẢN ĐẠI HỌC QUỐC GIA
TP HỒ CHÍ MINH - 2008

LỜI NÓI ĐẦU

Giáo trình này là một trong các giáo trình chính yếu của chuyên ngành Công nghệ thông tin. Giáo trình được xây dựng theo phương châm vừa đáp ứng yêu cầu chuẩn mực của sách giáo khoa, vừa có giá trị thực tiễn, đồng thời tăng cường khả năng tự học, tự nghiên cứu của sinh viên. Trên cơ sở đó, chúng tôi đã tham khảo nhiều tài liệu có giá trị của các tác giả trong và ngoài nước và đã sử dụng nhiều ví dụ lấy từ các ứng dụng thực tiễn.

Giáo trình này được dùng kèm giáo trình điện tử trên đĩa CD trong đó có thêm phần trình bày của giảng viên , các bài tập và phần đọc thêm nhằm đáp ứng tốt nhất cho việc tự học của sinh viên.

Chúng tôi rất mong nhận được các ý kiến đóng góp để giáo trình ngày càng hoàn thiện.

Nhóm biên soạn

MỤC LỤC

CHƯƠNG 1: KHÁI QUÁT VỀ XỬ LÝ ẢNH 11
1.1. Giới thiệu chung 11
1.2. Các bước cơ bản trong xử lý ảnh 12
1.3. Tài nguyên học tập 18
CHƯƠNG 2: ẢNH SỐ HÓA & SỰ CẢM NHẬN THỊ GIÁC ... 21
2.1. Các giá trị thông dụng 22
2.2. Những đặc trưng của các phép toán trên ảnh 23
2.2.1. Kiểu phép toán trên ảnh 23
2.2.2. Kiểu lân cận 24
2.3. Các tham số video 26
2.4. Sự cảm nhận tri giác 27
2.5. Độ nhạy sáng 28
2.5.1. Độ nhạy với bước sóng 28
2.5.2. Độ nhạy với tác nhân kích thích 28
2.6. Độ nhạy với tần số không gian 30
2.7. Độ nhạy về màu sắc 31
2.7.1. Bộ quan sát chuẩn 31
2.7.2. Các tọa độ màu CIE 32
2.8. Các áo giác quang học 35
CHƯƠNG 3: LẤY MẪU ẢNH 37
3.1. Độ trù mật lấy mẫu trong xử lý ảnh 38
3.1.1. Cửa chập lấy mẫu 38
3.2. Độ trù mật lấy mẫu trong phân tích ảnh 40
3.2.1. Lấy mẫu để đo diện tích 40
3.2.2. Lấy mẫu để đo độ dài 41
3.2.3. Kết luận về lấy mẫu 42

CHƯƠNG 4: CÁC LOẠI NHIỄU 44
4.1. Nhiễu photon 44
4.2. Nhiễu nhiệt 46
4.3. Nhiễu điện từ trên chip 47
4.4. Nhiễu KTC 47
4.5. Nhiễu khuếch đại 48
4.6. Nhiễu lượng hóa 18
CHƯƠNG 5: CÁC THIẾT BỊ THU ẢNH VÀ HIỂN THỊ ẢNH ... 50
5.1. Tính tuyến tính 50
5.2. Độ nhạy 51
5.2.1. Độ nhạy tuyệt đối 51
5.2.2. Độ nhạy tương đối 52
5.3. SNR 54
5.3.1. Nhiễu nhiệt (dòng tối) 54
5.3.2. Nhiễu photon 55
5.4. Tạo bóng 56
5.5. Dạng pixel 56
5.5.1. Các pixel vuông 57
5.5.2. Hệ số làm đầy 58
5.6. Độ nhạy phổ 59
5.7. Tốc độ chớp (thời gian tích hợp) 60
5.7.1. Tốc độ chớp trên các video camera 60
5.7.2. Tốc độ chớp trên các camera khoa học 60
5.8. Tốc độ đọc ra 61
5.9. Các thiết bị hiển thị ảnh 62
5.9.1. Tốc độ làm tươi 62
5.9.2. Hệ số đan xen 62
5.9.3 Độ phân giải 63

CHƯƠNG 6: CÁC PHÉP TOÁN XỬ LÝ ẢNH 64
6.1. Phép toán chập 64
6.2. Các tính chất của phép toán chập 64
6.3. Các biến đổi Fourier 65
6.4. Các tính chất của biến đổi Fourier 66
6.4.1. Tầm quan trọng của độ lớn và pha 69
6.4.2. Các tín hiệu đối xứng vòng 71
6.4.3. Ví dụ về các tính hiệu 2 – chiều phổ biến cùng với các biến đổi tương ứng của chúng 72
6.5. Các mô tả thống kê 75
6.5.1. Hàm phân phối xác suất về độ sáng 75
6.5.2. Hàm trù mật xác suất về độ sáng 76
6.5.3. Đại lượng trung bình 78
6.5.4. Độ lệch chuẩn 78
6.5.5. Hệ số biến đổi 79
6.5.6. Đại lượng phần trăm 79
6.5.7. Tỷ số tín hiệu – nhiễu 80
6.6. Biểu diễn biên của đối tượng 82
6.6.1. Mã xích 82
6.6.2. Các tính chất của mã xích 82
6.6.3. Mã crack 84
6.6.4. Mã chạy 85
6.7. Các phép toán trên lược đồ 85
6.7.1. Phép co giãn độ tương phản 85
6.7.2. Cân bằng lược đồ 86
6.7.3 Các phép toán dựa trên lược đồ khác 88
6.8. Các phép toán toán học 88
6.8.1. Các thao tác nhị phân 88
6.8.2. Các thao tác số học 91

6.9. Các phép toán dựa trên phép chập 91
6.9.1. Kiến thức nền 91
6.9.2. Phép chập trong miền tọa độ 94
6.9.3. Phép chập trong miền tần số 98
6.10. Các phép toán làm trơn ảnh 99
6.10.1. Các bộ lọc tuyến tính 99
6.10.2. Các bộ lọc phi tuyến 106
6.10.3. Tóm tắt các thuật toán làm trơn 108
6.11. Các phép toán đạo hàm 110
6.11.1. Các đạo hàm bậc nhất 111
6.11.2. Các đạo hàm bậc hai 116
6.11.3. Các bộ lọc khác 121
6.12. Các phép toán morphology 122
6.12.1. Các định nghĩa cơ bản 124
6.12.2. Phép phát triển và phép bào mòn 124
6.12.3. Phép chập logic 129
6.12.4. Phép mở và phép đóng 131
6.12.5. Phép trúng – trật 132
6.12.6. Tóm tắt những phép toán cơ bản 133
6.12.7. Phép rút xương 134
6.12.8. Phép lan truyền 138
6.12.9. Tóm tắt về phép rút xương & phép lan truyền .. 139
6.12.10. Xử lý giá trị xám theo kiểu morphology 141
6.12.11. Làm trơn theo kiểu morphology 144
6.12.12. Tính gradient theo kiểu morphology 144
6.12.13. Toán tử Laplace theo kiểu morphology 145
6.12.14. Tóm tắt về các bộ lọc morphology 145

CHƯƠNG 7: MỘT SỐ KỸ THUẬT HIỆU CHỈNH VÀ PHÂN ĐOẠN ẢNH 147
7.1. Kỹ thuật hiệu chỉnh sự tạo bóng 147
7.1.1. Mô hình tạo bóng 148
7.1.2. Ước lượng sự tạo bóng 148
7.2. Các kỹ thuật phục hồi và tăng cường ảnh cơ bản .. 153
7.2.1. Kỹ thuật tạo mặt nạ không sắc nét 154
7.2.2. Khử nhiễu 155
7.2.3. Khử độ méo của ảnh 157
7.3. Phân đoạn ảnh 160
7.3.1. Kỹ thuật lấy ngưỡng 160
7.3.2. Kỹ thuật dò cạnh 166
7.3.3. Kỹ thuật morphology nhị phân 173
7.3.4. Kỹ thuật morphology mức xám 179
TÀI LIỆU THAM KHẢO 183

Chương 1

KHÁI QUÁT VỀ XỬ LÝ ẢNH

1.1. GIỚI THIỆU CHUNG

Nhờ có công nghệ số hóa hiện đại, ngày nay con người đã có thể xử lý tín hiệu nhiều chiều thông qua nhiều hệ thống khác nhau, từ những mạch số đơn giản cho đến những máy tính song song cao cấp. Mục tiêu của việc xử lý này có thể chia làm ba hướng như sau:

- Xử lý ảnh ban đầu để có được ảnh mới theo một yêu cầu xác định (Ví dụ như ảnh mờ cần xử lý để được ảnh rõ hơn).
- Phân tích ảnh để thu được các thông tin đặc trưng giúp cho việc phân loại, nhận biết ảnh (Ví dụ: phân tích ảnh vân tay để trích chọn đặc trưng vân tay).
- Hiểu ảnh đầu vào để có những mô tả về ảnh ở mức cao hơn, sâu hơn (Ví dụ: từ ảnh một tai nạn giao thông phác họa hiện trường tai nạn).

Trong giáo trình này, chúng ta sẽ tập trung vào những khái niệm cơ bản của *xử lý ảnh* và giới hạn vấn đề trong phạm vi ảnh 2–chiều, mặc dù rằng nhiều khái niệm và kỹ thuật được trình bày ở đây hoàn toàn có thể mở rộng cho các ảnh số chiều cao hơn.

Chúng ta bắt đầu với những định nghĩa cơ bản nhất trong xử lý ảnh. Một ảnh trong thế giới được xem là một hàm

hai biến thực *a(x,y)* với a là độ sáng của ảnh tại vị trí tọa độ thực *(x,y)*. Một ảnh còn có thể chứa những ảnh con gọi là *các vùng quan tâm*, hay viết tắt theo tiếng Anh là *ROI (Region Of Interest)*. Các ROI này vẫn thường được gọi tắt là các *vùng*. Khái niệm vùng phản ánh một thực tế là trong ảnh thường chứa nhiều đối tượng, mỗi đối tượng tạo nên phần cơ sở của một vùng. Đối với một hệ xử lý ảnh cao cấp, chúng ta có thể áp dụng nhiều phép toán cho từng vùng ảnh một, ví dụ như một vùng ảnh này sẽ được áp dụng các phép toán bỏ hiệu ứng mờ do chuyển động, trong khi một vùng ảnh khác sẽ được xử lý để nâng cao chất lượng màu sắc của nó.

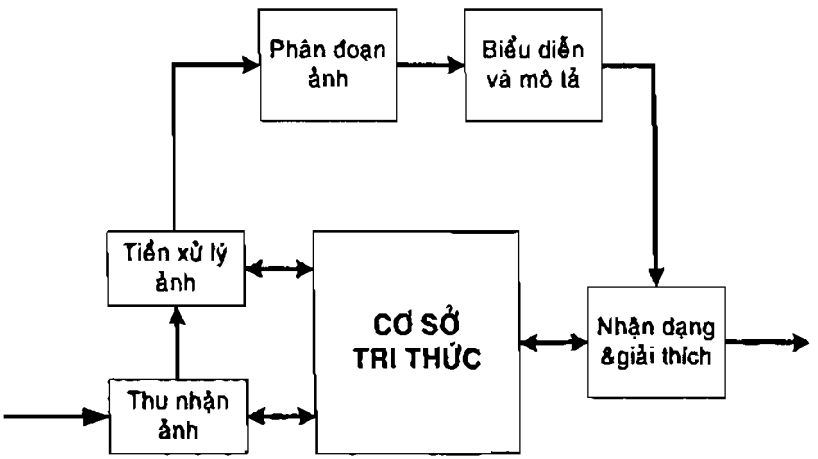
Các giá trị độ sáng của ảnh thường được thể hiện dưới dạng số thực hoặc số nguyên. Thông thường, những giá trị sáng kiểu số nguyên là kết quả của một quá trình lượng hóa chuyển một thang đo liên tục (chẳng hạn như giữa 0 và 100%) thành một số mức rời rạc. Tuy nhiên, trong nhiều quá trình hình thành ảnh, độ lớn của tín hiệu là kết quả đếm số hạt photon ở từng thời điểm, do vậy độ lớn ấy dĩ nhiên đã được lượng hoá sẵn. Còn trong một số quá trình tạo ảnh khác, ví dụ như tạo ảnh siêu âm trong y khoa, phương pháp đo đạc vật lý trực tiếp sẽ cho ra các giá trị phức, mỗi giá trị phức này gồm một giá trị độ lớn kết hợp với một giá trị pha, và cả hai đều ở dạng số thực.

1.2. CÁC BƯỚC CƠ BẢN TRONG XỬ LÝ ẢNH

Quá trình xử lý một ảnh số đầu vào nhằm thu được một ảnh đầu ra mong muốn thường phải trải qua nhiều bước khác nhau. Hình 1 giới thiệu với chúng ta những bước cơ bản của quá trình đó.

Trong hình vẽ này, để hoàn thành được mục tiêu xử lý cụ thể ứng với mỗi bước, chúng ta sẽ phải sử dụng, đến rất

nhiều khái niệm, định nghĩa, công cụ, thuật toán, kỹ thuật, vốn đã được phát triển và cải tiến trong quá trình hình thành nên ngành xử lý ảnh như ngày nay. Những kiến thức ấy sẽ lần lượt được trình bày trong những chương tiếp theo của giáo trình này.



Hình 1. *Các bước cơ bản trong xử lý ảnh*

Để hiểu rõ hơn về các bước đã được nêu trong hình 1, chúng ta hãy xét một ví dụ ứng dụng xử lý ảnh cụ thể. Giả sử là chúng ta cần xây dựng một chương trình cho phép xử lý từng bì thư một và qua đó cho ra địa chỉ của từng cặp người gửi/ nhận dưới dạng các chuỗi ký tự.

Với bài toán vừa nêu, rõ ràng việc trước hết cần phải thực hiện là có được ảnh bề mặt chứa địa chỉ của từng bì thư. Trong xử lý ảnh, công việc này có tên gọi là bước thu nhận ảnh, còn trong sơ đồ của hình 1 nó chính là bước đầu tiên. Để thực hiện, chúng ta cần có một *bộ cảm biến lấy ảnh* và khả

năng số hóa những tín hiệu liên tục được sinh ra bởi bộ cảm biến đó. Bộ cảm biến ở đây có thể là một máy chụp ảnh đơn sắc/ màu, hoặc một máy chụp ảnh kiểu quét dòng cho ra một dòng ảnh ở một thời điểm cụ thể. Với máy chụp kiểu này, sự di chuyển của bì thư sẽ giúp tạo ra một ảnh 2–chiều ứng với bề mặt của nó. Trong trường hợp kết xuất cho máy chụp ảnh/ bộ cảm biến cung cấp chưa phải ở dạng số hóa, chúng ta còn phải cho kết xuất đó đi qua một bộ phận khác nữa gọi là *bộ chuyển đổi tương tự – số,* vẫn hay được gọi tắt là ADC *(Analog–to–Digital Converter),* để thu được dạng số hóa của ảnh. Mặc dù đây chỉ là công đoạn đầu tiên trong quá trình xử lý, kết quả của nó có thể ảnh hưởng rất nhiều đến những công đoạn kế tiếp tùy theo loại hình ứng dụng, chất lượng, và chủng loại của thiết bị lấy ảnh. Những kiến thức có liên quan đến bước xử lý này sẽ lần lượt được khảo sát chi tiết hơn nữa trong các chương 2, 3, 5 của giáo trình.

Sau khi đã có ảnh bề mặt của từng bì thư dưới dạng số hóa, công đoạn kế tiếp là **tiền xử lý** ảnh đó. Ở bước này, ảnh sẽ được cải thiện về độ tương phản, khử nhiễu, khử bóng, khử độ lệch,... với mục đích làm cho chất lượng ảnh trở nên tốt hơn nữa, chuẩn bị cho các bước xử lý phức tạp hơn về sau trong quá trình xử lý ảnh. Các chương 4, 5, 6 và các mục 7.1, 7.2 sẽ cung cấp cho chúng ta những kiến thức có liên quan đến công đoạn tiền xử lý ảnh.

Bước thứ ba trong quá trình xử lý ảnh thường là bước **phân đoạn ảnh**. Nói một cách dễ hiểu, phân đoạn ảnh có nghĩa là chia một ảnh đầu vào thành nhiều phần nhỏ khác nhau hay còn gọi là các đối tượng. Trông có vẻ đơn giản, nhưng thật ra việc phân đoạn tự động ảnh thành tập những đối tượng khác nhau là nhiệm vụ phức tạp nhất trong xử lý

ảnh số hóa. Nếu kết quả phân đoạn ảnh chỉ dừng lại được ở mức thô thiển, thì toàn bộ những bước xử lý tiếp sau sẽ không cho kết quả tốt, và như vậy không thể đạt được thành công nào về mặt ứng dụng. Mặt khác, các thuật toán phân đoạn không đủ mạnh, hoạt động không ổn định cũng sẽ là nguồn gốc dẫn đến sự thất bại của một giải pháp xử lý ảnh. Trong bài toán đọc địa chỉ thư tín tự động đang xem xét, bước này đồng nghĩa với việc tách ra được từ nền ảnh những ký tự hay từ đã được biết trước là phần lớn thuộc về các "đối tượng địa chỉ".

Kết quả của bước phân đoạn ảnh vừa được đề cập trên đây thường được cho dưới dạng dữ liệu điểm ảnh thô, trong đó hàm chứa biên của một vùng ảnh, hoặc tập hợp tất cả những điểm ảnh thuộc về chính vùng ảnh đó. Trong cả hai trường hợp, sự chuyển đổi dữ liệu thô này thành một dạng thích hợp hơn cho việc xử lý trong máy tính là hết sức cần thiết. Để chuyển đổi chúng, câu hỏi đầu tiên cần phải trả lời là nên biểu diễn một vùng ảnh dưới dạng biên hay dưới dạng một vùng hoàn chỉnh gồm tất cả những điểm ảnh thuộc về nó. Biểu diễn dạng biên cho một vùng phù hợp với những ứng dụng chỉ quan tâm chủ yếu đến các đặc trưng hình dạng bên ngoài của đối tượng, ví dụ như các góc cạnh và điểm uốn trên biên chẳng hạn. Biểu diễn dạng vùng lại thích hợp cho những ứng dụng khai thác các tính chất bên trong của đối tượng, ví dụ như vân ảnh hoặc cấu trúc xương của nó. Trong một số ứng dụng, cả hai cách biểu diễn biên và vùng đều cần thiết, ví dụ như trong bài toán đọc thư mà chúng ta đang quan tâm. Bài toán này là một ví dụ điển hình cho lớp những ứng dụng nhận dạng ký tự, vốn luôn cần đến các thuật toán làm việc trên hình dạng biên cũng như trên cấu trúc xương, cùng nhiều tính chất bên trong của đối tượng.

Sự chọn lựa cách biểu diễn thích hợp cho một vùng ảnh chỉ mới là một phần trong việc chuyển đổi từ dữ liệu ảnh thô sang một dạng thích hợp hơn cho các xử lý về sau. Chúng ta còn phải đưa ra một phương pháp mô tả dữ liệu đã được chuyển đổi đó sao cho những tính chất cần quan tâm đến sẽ được làm nổi bật lên, thuận tiện cho việc xử lý chúng. Trong xử lý ảnh, thuật ngữ **mô tả** hay **chọn đặc trưng** đề cập đến sự rút trích từ ảnh những đặc trưng cần thiết dẫn đến sự hình thành các thông tin định lượng này có thể giúp chúng ta phân biệt các lớp đối tượng khác nhau trong ảnh. Như trong bài toán đọc thư nói riêng và nhận dạng ký tự nói chung, các bộ mô tả như "hồ" ("lỗ"), "vịnh" là những đặc trưng rất mạnh giúp phân biệt các phần khác nhau của bảng ký tự cần nhận dạng.

Bước cuối cùng trong sơ đồ đã nêu ra ở hình 1 là **nhận dạng** và **giải thích**. Nhận dạng là công đoạn gán *nhãn* cho một đối tượng dựa trên thông tin do bộ mô tả của đối tượng đó cung cấp. Giải thích là công đoạn gán nghĩa cho một tập các đối tượng đã được nhận biết. Trong ví dụ bài toán đọc thư đang xét, việc nhận dạng ra một ký tự như chữ "c" chẳng hạn đòi hỏi sự liên kết giữa các bộ mô tả cho ký tự đó với nhãn "c". Trong khi đó, việc giải thích thường là một bước khó hơn nữa so với nhận dạng, ví dụ như gán cho tập hợp gồm một dãy liên tiếp ba con số, theo sau là dấu "–", rồi lại theo sau đó là dãy liên tiếp hai con số, dấu "–", rồi lại đến một dãy liên tiếp bảy con số nữa, tức là chuỗi có dạng *xxx–xx–xxxxxxx*, một nghĩa là *"số điện thoại đầy đủ"*, bao gồm cả mã nước, mã vùng, và số điện thoại của từng nhà.

Cho đến đây chúng ta vẫn chưa để cập gì đến sự cần thiết của *tri thức cho trước* hay mối tương tác giữa *cơ sở tri thức* và các bước xử lý, như có thể quan sát thấy trong sơ đồ ở

hình 1. Ở đây, tri thức về một không gian bài toán sẽ được mã hóa vào trong một hệ xử lý ảnh dưới dạng cơ sở dữ liệu tri thức. Tri thức được đề cập đến có thể chỉ đơn giản là sự chi tiết hóa các vùng trong ảnh, nơi được biết trước là sẽ có những thông tin đáng quan tâm để tìm ra lời giải cho bài toán, do vậy giúp hệ thống thu hẹp phạm vi tìm kiếm và dẫn đến việc tìm ra lời giải nhanh hơn. Trong bài toán đọc thư, các vùng đó chính là những vị trí trên bì thư nơi thường có sự xuất hiện của các dòng địa chỉ liên quan đến người gửi/ nhận, tức là phần trái trên và phải dưới của ảnh bì thư đã thu được. Tri thức biết trước ấy cũng có thể hết sức phức tạp, ví dụ như bao gồm một danh sách liệt kê tất cả các khiếm khuyết chính có thể xảy ra ứng với bài toán kiểm định vật liệu, hoặc cả một cơ sở dữ liệu ảnh chứa nhiều ảnh chụp từ vệ tinh có độ phân giải cao cho một vùng như trong những ứng dụng kiểu dò tìm sự thay đổi trên mặt đất (theo dõi sự hình thành và phát triển của một cơn bão, của sự xây dựng một tổ hợp quân sự, ...). Ngoài mục đích hướng dẫn cách thức làm việc phù hợp cho mỗi bước xử lý ảnh, cơ sở tri thức còn có thể giúp điều khiển mối tương tác giữa các bước xử lý với nhau. Trong hình 1, sự tương tác này được thể hiện thông qua các mũi tên có hai đầu thay vì các mũi tên một đầu dùng để liên kết bước xử lý này với một bước xử lý tiếp sau. Sự mô tả như trong sơ đồ chỉ ra rằng các bước xử lý thường liên lạc với nhau trên cơ sở tri thức biết trước về dạng của kết quả mà chúng cần phải tạo ra. Lấy ví dụ, để nhận biết được chuỗi gồm tổng cộng 13 con số với hai dấu "–" nêu ở trên là một số điện thoại đầy đủ, thì hệ thống đọc thư tự động cần được cài đặt với tri thức giúp nhận ra vị trí đáng tin cậy của chuỗi ký tự đó khi nó thật sự là một số điện thoại, trong mối quan hệ với những thành phần khác trong vùng địa chỉ của bì thư. Bằng cách sử dụng

tri thức biết trước như vậy, chúng ta còn có thể nhận được sự phản hồi từ các bước xử lý thông qua cơ sở tri thức, chẳng hạn như chuỗi đó ở đúng vị trí của một số điện thoại đầy đủ, nhưng thông tin phản hồi cho biết nó chỉ gồm có 14 thay vì 15 ký tự, từ đó kết luận được là có thể đã có hai ký tự dính nhau. Với tình huống này, sự phản hồi thông qua cơ sở tri thức sẽ giúp hệ thống ra quyết định cho bước phân đoạn ảnh tiếp tục thực hiện lại việc phân đoạn một cách chính xác hơn. Có thể xem đây là một ví dụ về sự tận dụng tri thức trong khi thực hiện các tác vụ xử lý ảnh.

Trong những bước xử lý ảnh được thảo luận trên đây, chúng ta vẫn chưa bàn đến bài toán hiển thị ảnh, cũng là một bài toán hết sức phức tạp. Nó sẽ được trình bày sau chương 5 của giáo trình. Tuy nhiên, có một điều quan trọng mà chúng ta cần phải lưu ý là trong sơ đồ ở hình 1, yêu cầu hiển thị ảnh có thể được đưa ra ở bất kỳ một bước xử lý nào. Hơn nữa, không phải ứng dụng xử lý ảnh nào cũng cần phải thực hiện hết tất cả các bước xử lý đã nêu, ví dụ như các ứng dụng chỉnh sửa ảnh nghệ thuật chỉ dừng lại ở bước tiền xử lý. Một cách tổng quát thì những chức năng xử lý bao gồm cả nhận dạng và giải thích thường chỉ có mặt trong các hệ thống phân tích ảnh tự động hoặc bán tự động, được dùng để rút trích ra những thông tin quan trọng từ ảnh, ví dụ như các ứng dụng nhận dạng ký tự quang học.

1.3. TÀI NGUYÊN HỌC TẬP

Xử lý ảnh là một ngành nghiên cứu trẻ rất năng động trong những thập niên gần đây, do vậy có một khối lượng khá lớn các công trình nghiên cứu lý thuyết và ứng dụng đã được xuất bản. Phần lớn các công trình này có thể tìm thấy trên các

web site nghiên cứu của từng nhà nghiên cứu một, cũng như trong các web site thư viện điện tử của IEEE (www.ieee.org), IEEE Computer Society (www.computer.org/publications/), ACM (www.acm.org), Elsevier Science (www.elsevier.nl). Sau đây là danh sách các tạp chí và sách tham khảo chuyên ngành phổ biến nhất về xử lý ảnh.

Tạp chí nghiên cứu chuyên ngành

IEEE Computer Graphics & Applications

IEEE Transactions on Multimedia

IEEE Transactions on Pattern Analysis and Machine Intelligence

Sách tham khảo kinh điển

Tựa đề	Tác giả	Nhà xuất bản
Digital Picture Processing	A. Rosenfeld and A.C.Kak	Academic Press, New York, 1976 (lần 1), 1981 (lần 2)
Digital Image Processing	W.K.Pratt	Wiley, New York, 1978.
Fundamental of Digital Image Processing	Anil K.Jain	Prentice–Hall, Englewood Cliffs, New Yersey, 1989.
Rafael C.Gonzalez, Richard E.Woods	Digital Image Processing	Addison – Wesley, 1993

Bạn đọc cũng có thể tìm thấy rất nhiều nguồn tư liệu khác trên Internet dựa vào một số trang chủ chính, ví dụ như trang chủ sau đây: http://www.ee.princeton.edu/~icip95/iplink/. Ngoài ra, nếu bạn đọc muốn đi sâu vào nghiên cứu thì ứng với mỗi chủ đề được trình bày trong giáo trình này, bao giờ cũng có kèm theo rất nhiều các bài báo, công trình có liên quan được liệt kê trong phần Tài liệu tham khảo ở cuối sách.

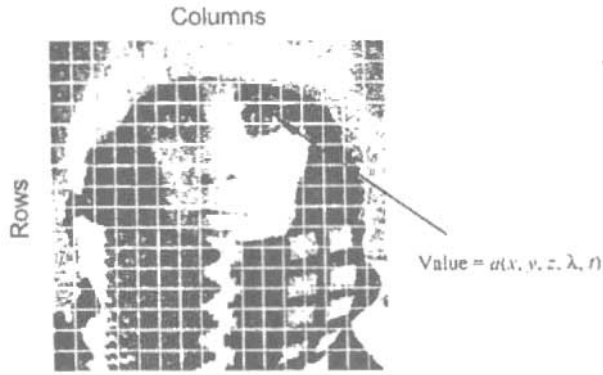
Chương 2

ẢNH SỐ HÓA VÀ SỰ CẢM NHẬN THỊ GIÁC

Một ảnh số *a[m,n]* được mô tả trong không gian rời rạc 2–chiều chính là kết quả của quá trình lấy mẫu ảnh ở dạng tương tự *a(x,y)* trong không gian liên tục 2–chiều. Quá trình này vẫn hay được đề cập đến với tên gọi là quá trình số hóa, với tác động của nó trên một ảnh liên tục được chỉ ra trong hình 2, còn bản chất toán học được khảo sát trong chương 3 của giáo trình.

Ảnh liên tục 2–chiều *a(x,y)* được chia thành *N hàng* và *M cột*. Phần giao của mỗi hàng và cột được gọi là một *pixel (điểm ảnh)*. Các giá trị độ sáng được gán cho các tọa độ nguyên [*m,n*], với {*m* = 0, 1, 2,... , *M*–1} và {*n* = 0, 1, 2..., *N*–1}, tạo thành một mảng *a[m,n]*. Trong thực tế thì gần như ở tất cả các trường hợp, *a(x, y)* – vốn có thể xem là một tín hiệu vật lý tiếp xúc với bề mặt của một bộ cảm biến 2–chiều – thật ra là một hàm nhiều biến, gồm các biến như **độ sâu (z), màu sắc (λ), thời gian (t).** Trong chương này chúng ta chỉ xét các ảnh tĩnh đơn sắc 2–chiều.

Ảnh trong hình 2 đã được chia thành $N = 16$ hàng và $M = 16$ cột. Giá trị được gán cho mỗi pixel chính là giá trị độ sáng trung bình của nó được làm tròn thành số nguyên gần nhất. Quá trình biểu diễn độ lớn của một tín hiệu 2–chiều tại một tọa độ cụ thể bằng một con số nguyên với L mức xám khác nhau thường được gọi là *quá trình lượng hóa.*



Hình 2. *Quá trình số hóa một ảnh liên tục. Pixel ở tọa độ [m = 10, n = 3] có giá trị độ sáng là số nguyên bằng 110.*

2.1. CÁC GIÁ TRỊ THÔNG DỤNG

Bảng 1 dưới đây liệt kê một số giá trị chuẩn cho ba tham số thông dụng nhất trong xử lý ảnh là hàng, cột, và các mức xám.

Bảng 1. *Các giá trị phổ biến trong xử lý ảnh*

Tham số	Ký hiệu	Các giá trị chuẩn
Hàng	*N*	256, 512, 525, 625, 1024, 1035
Cột	*M*	256, 512, 768, 1024. 1320
Các mức xám	*L*	2,64, 256, 1024, 4096, 16384

Chúng ta rất hay gặp trường hợp $M = N = 2^k$ với (K = 8, 9, 10) khi thực hiện một số thuật toán như biến đổi Fourier nhanh (xem mục 6.3 của Chương 6).

Bạn đọc cũng nên lưu ý là số các mức xám phân biệt thường ở dạng lũy thừa của 2, tức là $L = 2^B$, trong đó B là số bit dùng trong biểu diễn nhị phân của các giá trị độ sáng

ảnh. Khi B > 1 chúng ta có *ảnh mức xám*, còn khi $B = 1$ chúng ta có *ảnh nhị phân*. Ảnh kiểu này chỉ có hai mức xám là "đen" và "trắng", hay "0" và "1".

2.2 NHỮNG ĐẶC TRƯNG CỦA CÁC PHÉP TOÁN TRÊN ẢNH

Ứng với mỗi lớp ghép toán trên ảnh cho trước, người ta thường có hai mối quan tâm. Mối quan tâm thứ nhất là lớp phép toán ấy cho ra kết quả cụ thể như thế nào trên ảnh. Mối quan tâm còn lại là phí tổn tính toán cần có để thực hiện những phép toán ấy. Muốn trả lời hai câu hỏi quan trọng này, chúng ta cần phải tiến hành phân loại các phép toán trên ảnh.

2.2.1. Kiểu phép toán trên ảnh

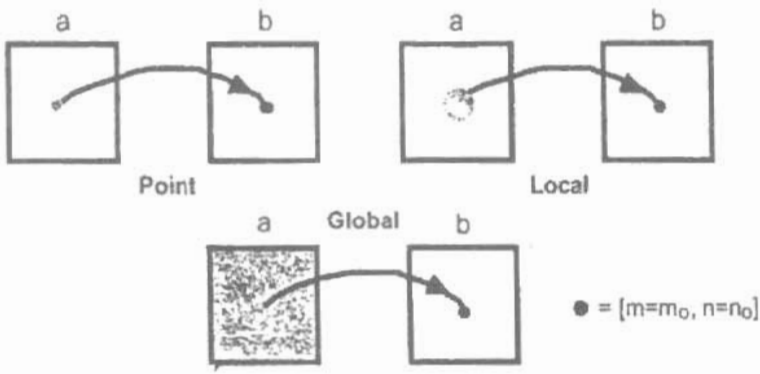
Một cách tổng quát, các phép toán trên ảnh cho phép biến đổi một ảnh đầu vào *a[m, n]* thành một ảnh đầu ra *b[m, n]* (hoặc thành một cách biểu diễn ảnh khác) có thể chia thành ba lớp như trong bảng 2 dưới đây.

Minh họa của các kiểu phép toán nói trên được cho trong hình sau đây.

Bảng 2. *Các loại phép toán trên ảnh. Kích thước ảnh ở đây là N x N, còn kích thước lân cận là P x P. Chúng ta chú ý rằng độ phức tạp được tính bằng số phép toán ứng với mỗi pixel.*

Phép toán	Tính năng	Độ phức tạo chung/ Pixel
Trên pixel	Giá trị đầu ra tại một tọa độ cụ thể chỉ phụ thuộc vào giá trị đầu vào tại cùng tọa độ đó.	*Hằng số*

Cục bộ	Giá trị đầu ra tại một tọa độ cụ thể phụ thuộc vào các giá trị đầu vào trong lân cận của tọa độ đó.	P^2
Toàn cục	Giá trị đầu ra tại một tọa độ cụ thể phụ thuộc vào tất cả các giá trị của ảnh đầu vào.	N^2



Hình 3. *Minh họa các kiểu phép toán trên ảnh*

2.2.2 Kiểu lân cận

Trong xử lý ảnh, những phép toán dựa trên lân cận giữ vai trò hết sức quan trọng. Chính vì lý do đó, chúng ta cần tìm hiểu cách thức ảnh được lấy mẫu, cũng như mối quan hệ giữa cách thức đó với rất nhiều kiểu lân cận có thể sử dụng trong quá trình xử lý một ảnh số.

Lấy mẫu theo hình chữ nhật

Trong hầu hết các trường hợp, ảnh sẽ được lấy mẫu bằng cách đặt lên trên nó một lưới chữ nhật, như đã chỉ ra trong

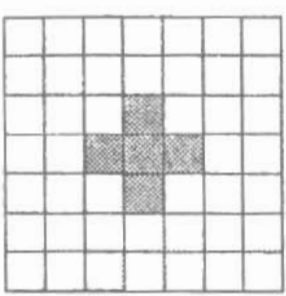
hình 2. Cách làm này dẫn đến kiểu lấy mẫu được nêu trong hình 4(a) và 4(b).

Lấy mẫu theo hình lục giác

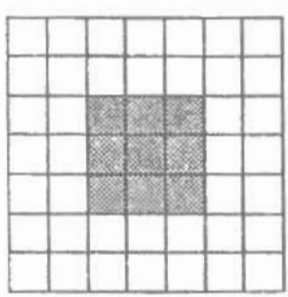
Một cách lấy mẫu khác được minh họa trong hình 4(c). Nó được gọi một cách ngắn gọn là lấy mẫu lục giác.

Cả hai cách lấy mẫu nói trên đều đã được nghiên cứu khá nhiều [1] và điều có khả năng lấp đầy không gian ảnh liên tục theo kiểu lặp đi lặp lại. Tuy nhiên do những ràng buộc về phần cứng và phần mềm, cách lấy mẫu theo hình chữ nhật được sử dụng phổ biến hơn.

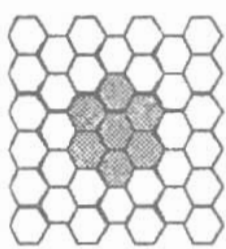
Các phép toán cục bộ cho ra một giá trị kết xuất $b[m = m_0, n = n_0]$ dựa trên giá trị của những pixel nằm trong *lân cận* của pixel $b[m = m_0, n = n_0]$. Trong hình 4 dưới đây, chúng ta có thể thấy một số lân cận phổ biến, ví dụ như lân cận liên thông 4 và lân cận liên thông 8 (theo cách lấy mẫu hình chữ nhật), lân cận liên thông 6 (theo cách lấy mẫu lục giác).



Hình 4(a)
Lấy mẫu hình chữ nhật liên thông 4



Hình 4(b)
Lấy mẫu hình chữ nhật liên thông 8



Hình 4(c)
Lấy mẫu lục giác liên thông 6

2.3. CÁC THAM SỐ VIDEO

Do có nhiều ảnh tĩnh được thu nhận từ các máy quay video và các máy bắt khung, chúng ta cũng cần đề cập đến các tiêu chuẩn có liên quan đến ba dạng biểu diễn tín hiệu video phổ biến hiện nay trên thế giới là các hệ phát hình NTSC, PAL và SECAM. Thông tin về các chuẩn này được cho trong bảng 3 sau đây.

Bảng 3. *Các tham số video chuẩn*

Chuẩn	NTSC	PAL	SECAM
Số ảnh/giây	29.97	25	25
Lượng miligiây/ ảnh	33.37	40.0	40.0
Số dòng/ ảnh	525	625	625
(ngang/ dọc) = hệ số tỷ lệ	4:3	4:3	4:3
Kiểu đan xen	2:1	2:1	2:1

Trong một ảnh đan xen, các dòng số lẻ (1, 3, 5...) được quét ở nửa đầu của khoảng thời gian được cấp phát (tức là 20 ms trong hệ PAL), còn các dòng số lẻ được quét ở nửa sau (2, 4, 6...). Màn hình hiển thị ảnh khi đó cần phải được chia tọa độ theo định dạng quét này (xem mục 5.9.2). Nguyên nhân chính của việc tạo sự đan xen trong các dòng quét của một ảnh video là nhằm giảm độ rung của ảnh khi được hiển thị trên màn hình. Nếu chúng ta có ý định sử dụng các ảnh được lấy vào từ một nguồn video phát theo kiểu đan xen, điều quan trọng là chúng ta cần phải chú ý xem hai nửa ảnh có được "xê dịch" một cách phù hợp không thông qua phần cứng số hóa, hay là nó nên được cài đặt bằng phần mềm. Ngoài ra, việc

phân tích những đối tượng chuyển động đòi hỏi những chú ý đặc biệt đến nguồn video đan xen nhằm tránh được các cạnh chạy theo hình chữ chi.

Số hàng N của một nguồn phát video thường tương ứng một – một với số dòng của một ảnh video. Trong khi đó, số cột lại phụ thuộc vào cơ chế điện tử được dùng để số hóa ảnh. Các máy bắt khung khác nhau có thể sản sinh ra M = 384 hay 768 cột ứng với mỗi dòng cho cùng một nguồn video đầu vào.

2.4. SỰ CẢM NHẬN THỊ GIÁC

Ngày nay chúng ta có rất nhiều ứng dụng xử lý ảnh được thiết kế nhằm tạo ra các ảnh số phục vụ cho sự quan sát, thưởng thức của con người, thay vì phục vụ cho các quá trình giám định công nghiệp tự động như trong các nhà máy sản xuất có sự trợ giúp của máy tính. Khi thiết kế những ứng dụng như vậy, sự hiểu biết về các đặc trưng và giới hạn trong hệ thống thị giác của con người là hết sức quan trọng, vì chúng sẽ góp phần làm nên sự thành công của ứng dụng. Ít nhất chúng ta cũng phải ý thức được rằng:

(1) Chưa có sự hiểu biết rõ ràng về hệ thống thị giác của nhân loại.

(2) Hoàn toàn chưa cóp một độ đo nào có thể dùng để đánh giá chất lượng của hình ảnh dựa trên sự cảm nhận về chất lượng thị giác của con người.

(3) Không tồn tại một mô hình quan sát kiểu mẫu nào cho hệ thống thị giác của con người.

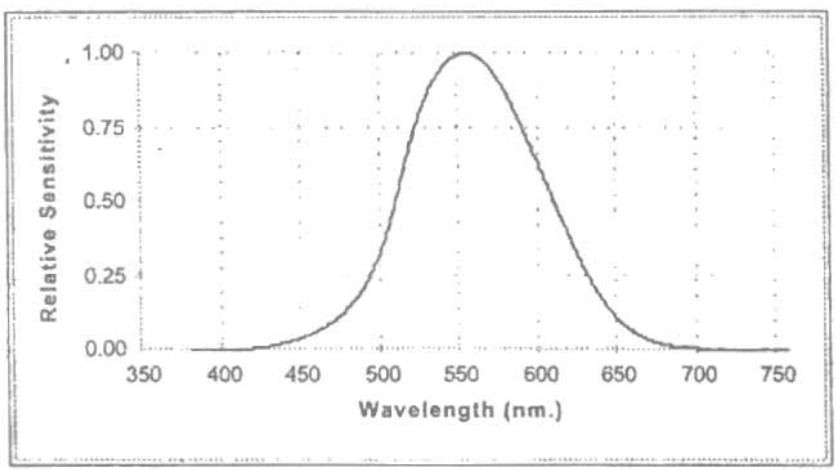
Mặc dù vậy, những nghiên cứu gần đây về tâm lý giác quan cũng đã mang lại cho con người những hiểu biết sâu sắc hơn về hệ thống thị giác của chính mình (xem Stockham [12]).

2.5. ĐỘ NHẠY SÁNG

Có rất nhiều cách mô tả độ nhạy trong hệ thống thị giác của con người. Chúng ta giả sử là một vùng đồng nhất trong một ảnh số sẽ có cường độ sáng cho bởi một hàm theo bước sóng (màu sắc) $I(\lambda)$. Ngoài ra $I(\lambda) = I_0$ là một giá trị không đổi.

2.5.1. Độ nhạy với bước sóng

Hình 5 dưới đây chỉ ra cường độ cảm nhận, hay còn gọi là độ nhạy phổ, dưới dạng một hàm theo bước sóng λ.



Hình 5. *Độ nhạy phổ ứng với một bộ quan sát kiểu con người*

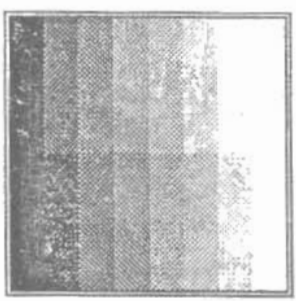
2.5.2. Độ nhạy với tác nhân kích thích

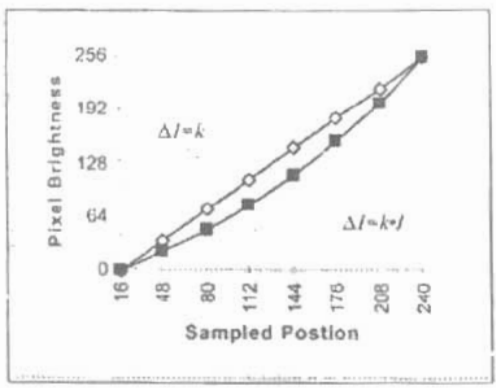
Nếu cường độ sáng không đổi I_0 được cho phép thay đổi thì với một xấp xỉ khá tốt, chúng có đáp ứng thị giác R sẽ tỷ lệ với logarit của cường độ đó. Kết quả này gọi là luật Weber–Fechner:

$$R = \log(I_0) \qquad (1)$$

Những ý tưởng xoay quanh luật vật lý này rất dễ minh họa. Ở đây các bước cảm nhận bằng nhau về độ sáng $\Delta R = k$ đòi hỏi độ sáng vật lý (tác nhân kích thích) phải tăng theo lũy thừa. Điều này được chỉ ra trên hình 6(a) và 6(b).

Một đường nằm ngang đi qua phần đỉnh của hình 6(a) chỉ ra có sự gia tăng tuyến tính về độ sáng khách quan (hình 6(b)), nhưng độ sáng chủ quan lại tăng theo hàm logarit. Một đường nằm ngang khác đi qua phần đáy của hình 6(a) thể hiện sự gia tăng lũy thừa trong độ sáng khách quan (hình 6(b)), nhưng độ sáng chủ quan lại tăng tuyến tính.





Hình 6(a)
(đỉnh) Bước sáng –I = k
(đáy) Bước sáng –I = k.I

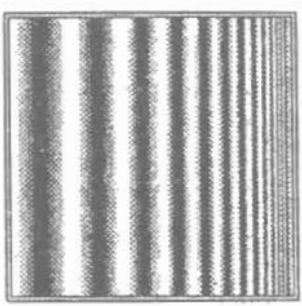
Hình 6(b)
Các độ sáng thật sự cộng thêm các giá trị nội suy

Trong hình 6 (a) chúng ta có thể quan sát thấy hiệu ứng dải tần Mach. Mặc dù độ sáng vật lý là không đổi khi chúng ta đi qua các dải dọc, bộ quan sát của con người khi đó sẽ cảm nhận ra có một lượng thừa và thiếu về độ sáng tại vị trí cạnh của mỗi bước sáng. Do vậy ngay trước khi sang một bước sáng

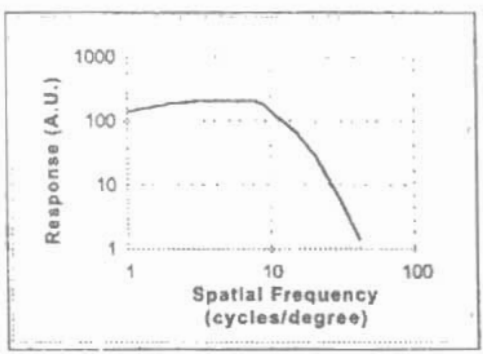
mới, chúng ta thấy có một sự giảm nhẹ về độ sáng so với giá trị vật lý thật sự. Còn ngay sau khi sang bước sáng mới, chúng ta lại thấy có một sự tăng nhẹ về độ sáng so với giá trị vật lý thật sự. Hiệu ứng tổng thể là sự gia tăng độ tương phản cảm nhận cục bộ tại vị trí cạnh của mỗi bước sáng.

2.6. ĐỘ NHẠY VỚI TẦN SỐ KHÔNG GIAN

Nếu cường độ sáng không đổi I_0 được thay thế bằng một sóng hình sin tăng dần về tần số không gian (hình 7(a)), chúng ta có thể đi xác định độ nhạy với tần số không gian như được chỉ ra trong hình 7 (b) [14, 15].



Hình 7(a). *Sóng kiểm tra hình sin*



Hình 7(b). *Độ nhạy với tần số không gian*

Để diễn dịch lại các dữ liệu này sang dạng thuật ngữ thông thường, chúng ta hãy xem một màn hình máy tính lý tưởng với khoảng cách quan sát là 50cm. Tần số không gian cho đáp ứng cực đại là 10 chu kỳ/ độ (xem hình 7(b)). Một độ quan sát ở mức 50cm sẽ được chuyển đổi thành $50 \times \tan(1^0) = 0.87cm$ trên màn hình máy tính. Do vậy tần số không gian

của đáp ứng cực đại f_{max} = *10 chu kỳ/0.87 cm = 11.46 chu kỳ/cm* tại khoảng cách quan sát này. Biến đổi công thức này thành dạng tổng quát hơn chúng ta có:

$$f_{max} = \frac{10}{d \tan(1^0)} \text{ chu kỳ/cm} \qquad (2)$$

trong đó d là khoảng cách quan sát được đo theo đơn vị cm.

2.7. ĐỘ NHẠY VỀ MÀU SẮC

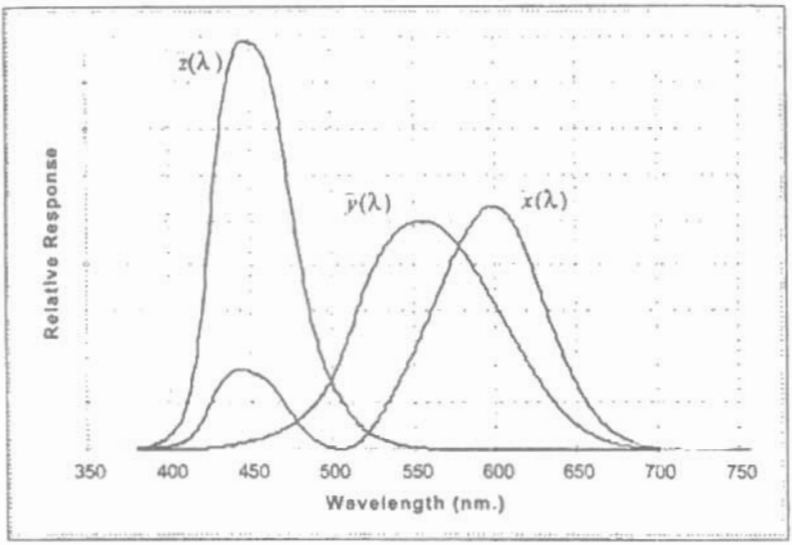
Sự cảm nhận của con người về màu sắc là một chủ đề vô cùng phức tạp. Do vậy trong mục này chúng ta chỉ trình bày tóm tắt một số kết quả quan trọng nhất. Như đã biết, sự cảm nhận màu sắc về mặt vật lý được dựa trên ba sắc màu có trong võng mạc của mắt người.

2.7.1 Bộ quan sát chuẩn

Dựa trên các phép đo tâm vật lý, Ủy ban CIE (Commission Internationale de I'Elairage) đã thừa nhận một số đường cong chuẩn làm đường cong độ nhạy cho một bộ quan sát kiểu mẫu ứng với ba "sắc màu" $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$, như được thể hiện trong hình 8. Đây không phải là các đặc trưng thu hút màu sắc thật sự trong một võng mạc chuẩn của con người, mà chỉ đơn thuần là các đường cong độ nhạy có được từ dữ liệu thực tế [10].

Trong một ảnh, với mỗi vùng đồng nhất có cường độ sáng là hàm theo bước sóng (màu sắc) $I(\lambda)$, chúng ta có ba đáp ứng:

$$X = \int_0^{\infty} I(\lambda)\bar{x}(\lambda)d\lambda \quad Y = \int_0^{\infty} I(\lambda)\bar{y}(\lambda)d\lambda \quad Z = \int_0^{\infty} I(\lambda)\bar{z}(\lambda)d\lambda \qquad (3)$$



Hình 8. *Các đường cong độ nhạy màu ứng với một bộ quan sát chuẩn*

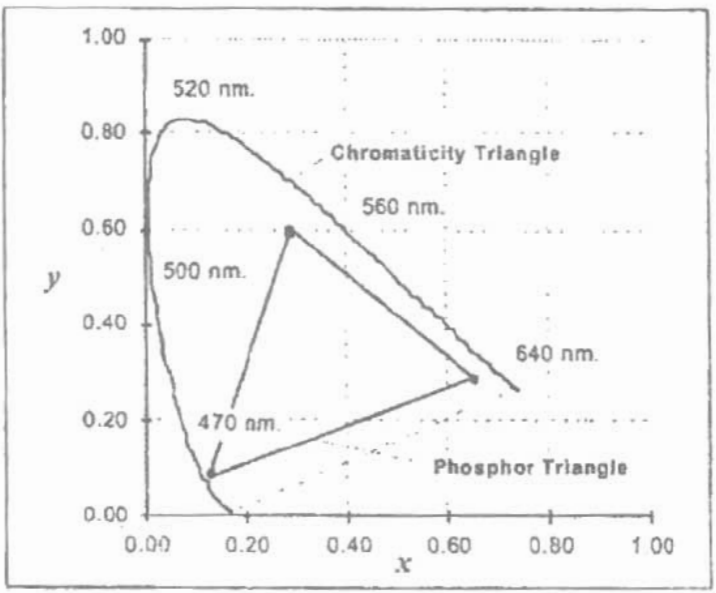
2.7.2. Các tọa độ màu CIE

Các tọa độ màu mô tả lượng thông tin màu sắc cảm nhận được bởi mắt người. Chúng được xác định thông qua các công thức sau đây:

$$x = \frac{X}{X+Y+Z} \qquad y = \frac{Y}{X+Y+Z} \qquad z = 1 - (x + y) \quad (4)$$

Trong các công thức trên, x đại diện cho tọa độ màu đỏ, còn y đại diện cho tọa độ màu lục. Các giá trị này tuyến tính trong $I(\lambda)$, do đó thông tin về cường độ tuyệt đối đã bị loại mất trong phép tính tọa độ màu (x,y) mà chúng ta vừa nêu. Toàn bộ các phân phối màu $I(\lambda)$, xuất hiện trước bộ quan sát với cùng một sắc màu sẽ có cùng tọa độ màu sắc theo cách tính này.

Nếu chúng ta sử dụng một nguồn màu thuần tinh chỉnh được, thì cường độ có thể được mô hình hóa dưới dạng $I(\lambda) = \delta(\lambda - \lambda_0)$ với $\delta(\bullet)$ là hàm xung. Tập hợp các tọa độ màu (x,y) sinh bởi sự biến thiên của λ_0 được gọi là tam giác màu CIE, như được minh họa trong hình 9.



Hình 9. *Lược đồ màu chứa tam giác màu CIE liên kết với các màu có phổ thuần và tam giác ứng với các chất phát sáng trên màn hình CRT*

Trong hình 9, các màu có phổ thuần sẽ nằm dọc theo biên của tam giác màu. Tất cả những màu còn lại nằm bên trong của tam giác. Bảng 4 dưới đây cho chúng ta tọa độ màu của một số nguồn sáng tiêu biểu.

Bảng 4. *Tọa độ màu của một số nguồn sáng chuẩn*

Nguồn	x	y
Đèn huỳnh quang với 4800° K	0.35	0.37
Mặt trời với 6000° K	0.32	0.33
Chất phát sáng đỏ (europium yttrium vanadate)	0.68	0.32
Chất phát sáng lục (zinc cadmium sulfide)	0.28	0.60
Chất phát sáng lam (zinc sulfide)	0.15	0.07

Sự mô tả màu sắc dựa trên các tọa độ màu không những cho phép phân tích màu còn giúp tổng hợp được chúng. Bằng cách trộn hai nguồn màu lại với nhau, chúng ta có thể tạo ra một màu bất kỳ nằm dọc theo đường thẳng nối các tọa độ màu tương ứng với hai màu đó. Do số lượng hạt photon không thể có giá trị âm nên các hệ số trộn phải dương. Bằng cách sử dụng ba nguồn màu như các chất phát sáng đỏ, lục, lam trên các màn hình CRT, chúng ta sẽ có được tập hợp màu được xác định bởi phần bên trong của "tam giác phát sáng" như trong hình 9.

Công thức để chuyển đổi từ các giá trị phát sáng (X, Y, Z) sang các màu màn hình CRT thông dụng (R, G, B) và ngược lại được cho trong các công thức sau đây:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.9107 & -0.5326 & -0.2883 \\ -0.9843 & 1.9984 & -0.0283 \\ 0.0583 & -0.1185 & 0.8986 \end{bmatrix} \qquad (5)$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.6067 & 0.1736 & 0.2001 \\ 0.2988 & 0.5868 & 0.1143 \\ 0.0000 & 0.0661 & 1.1149 \end{bmatrix} \bullet \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad (6)$$

Khi vị trí của một màu mong muốn *(X, Y, Z)* ở trong tam giác phát sáng của hình 9, các giá trị R, G, B tính được nhờ phương trình (5) luôn dương và do vậy có thể dùng để điều khiển sự phát màu trên màn hình CRT.

Ý tưởng cho rằng ứng với một độ dịch chuyển nhỏ trong lược đồ màu (hình 9), chúng ta sẽ thu được một cách tỷ lệ một lượng thay đổi nhỏ về màu sắc cảm nhận là không chính xác. Tính chất đó chỉ được xấp xỉ trong không gian màu có được từ thực nghiệm) – sau đây:

$$u' = \frac{4x}{-2x + 12y + 3} \qquad v' = \frac{9y}{-2x + 12y + 3}$$

$$x = \frac{9u'}{6u' - 16v' + 12} \qquad y = \frac{4v'}{6u' - 16v' + 12} \qquad (7)$$

2.8. CÁC ẢO GIÁC QUANG HỌC

Sự mô tả về hệ thống thị giác của con người mà chúng ta vừa trình bày ở trên hoàn toàn được thể hiện thông qua các thuật ngữ kỹ thuật chuẩn mực. Điều này dẫn đến một kết luận là, chúng ta có đủ trí thức về hệ thống thị giác của con người để có thể mô hình hóa nó theo những kỹ thuật phân tích hệ thống tiêu biểu. Tuy nhiên, hai ví dụ đơn giản về ảo giác quang học mà chúng ta chỉ ra sau đây trên màn hình 10 cho thấy rằng cách tiếp cận có tính hệ thống như vậy đã đơn giản hóa vấn đề đi rất nhiều. Do vậy, khi sử dụng những mô hình dại loại như vậy, chúng ta phải hết sức cẩn thận.

Hình 10. *Các ảo giác quang học*

Ảnh bên trái trong hình 10 thể hiện ảo giác về các giá trị xám trong mắt người khi bộ não biết rằng thực tế chúng không tồn tại. Hơn thế nữa, chúng ta có cảm giác về một sự thay đổi động nào đó trong ảnh, mà một phần là do các chuyển động của mắt người. Ảnh bên phải trong hình 10, được gọi là tam giác Kanizsa, minh họa cho sự tăng cường về một tương phản và sự xuất hiện của các đường biên giả tạo [14]. Chúng ta sẽ không thể giải thích được những hiện tượng này dựa vào hệ thống cảm nhận thị giác mà chúng ta đã trình bày theo định hướng hệ thống ở trên.

Chương 3

LẤY MẪU ẢNH

Quá trình lấy mẫu ảnh là quá trình chuyển đổi một ảnh liên tục *a(x, y)* sang dạng biểu diễn số hóa *b[m, n]*. Trong một hệ thống lấy mẫu lý tưởng, *a(x, y)* sẽ được nhân với một tín hiệu huấn luyện xung lý tưởng 2–chiều:

$$b_{ideal}[m,n] = a(x,y) \bullet \sum_{m=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \delta(x - mX_0, y - nY_0)$$

$$= \sum_{m=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} a(mX_0, nY_0)\delta(x - mX_0, y - mX_0) \quad (8)$$

với X_0 và Y_0 là các khoảng lấy mẫu, $\delta(\bullet, \bullet)$ là một hàm xung lý tưởng. Ở một điểm nào đó dĩ nhiên hàm xung $\delta(x, y)$ sẽ được chuyển đổi thành hàm xung rời rạc $\delta[m, n]$. Với cách lấy mẫu theo lưới vuông, chúng ta có $X_0 = Y_0$. Việc lấy mẫu thông qua một hàm xung tương ứng với sự lấy mẫu trên các điểm vô cùng nhỏ. Tuy nhiên cách lấy mẫu này lại không giống với tình huống thông thường trong hình 2. Để thể hiện tác động của một cửa chập lấy mẫu hữu hạn p(x, y), chúng ta có thể chỉnh sửa mô hình lấy mẫu lại như sau:

$$b[m,n] = (a(x,y) \otimes p(x,y)) \bullet \sum_{m=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \delta(x - mX_0, y - nY_0) \quad (9)$$

Tác động kết hợp của cửa chập và sự lấy mẫu d89ược hiểu rõ ràng nhất khi chúng ta xét biểu diễn của tín hiệu trên miền Fouier.

$$B[\Omega, \Psi] = \frac{1}{4\pi^2} \sum_{m=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} A(\Omega - m\Omega_s, \Psi - n\Psi_s) \bullet P(\Omega - m\Omega_s, \Psi - n\Psi_s) \quad (10)$$

trong đó $\Omega_x = 2\pi/X_0$ là tần số lấy mẫu theo hướng x, $\Psi_x = 2\pi/Y_0$ là tần số lấy mẫu theo hướng y. Cửa chập $p(x, y)$ thường có dạng vuông, tròn, hay Gauss với xác suất liên kết $P(\Omega, \Psi)$ (xem bảng 11). Bản chất tuần hoàn của phổ nói trên, như được mô tả trong phương trình (53), có thể được quan sát thấy rất rõ ràng từ phương trình (10).

3.1. ĐỘ TRÙ MẬT LẤY MẪU TRONG XỬ LÝ ẢNH

Để ngăn ngừa hiện tượng *aliasing* (chồng lấp lên nhau) của các số hạng phổ như có thể thấy trước được từ phương trình (10), hai điều kiện sau cần phải được thỏa mãn:

A(u,v) là tín hiệu có dải tần giới hạn –

$$|A(u,v) \equiv 0 \quad \text{với } |u| > uc \text{ và } |v| > v_c \quad (11)$$

Tần số lấy mẫu Nyquist –

$$\Omega_s > 2 \bullet u_c \quad \text{và} \quad \Psi s > 2 \bullet v_c \quad (12)$$

trong đó u_c và v_c là các tần số cắt theo hai hướng x và y tương ứng. Các ảnh được thu nhận từ những thấu kính có tính đối xứng vòng, không có quang sai, có độ nhiễu xạ giới hạn nhìn chung sẽ có dải tần giới hạn. Các thấu kính đó hoạt động như một bộ lọc thông thấp với tần số cắt trong miền tần số phương trình (43) được cho bởi:

$$u_c = v_c = \frac{2NA}{\lambda} \quad (13)$$

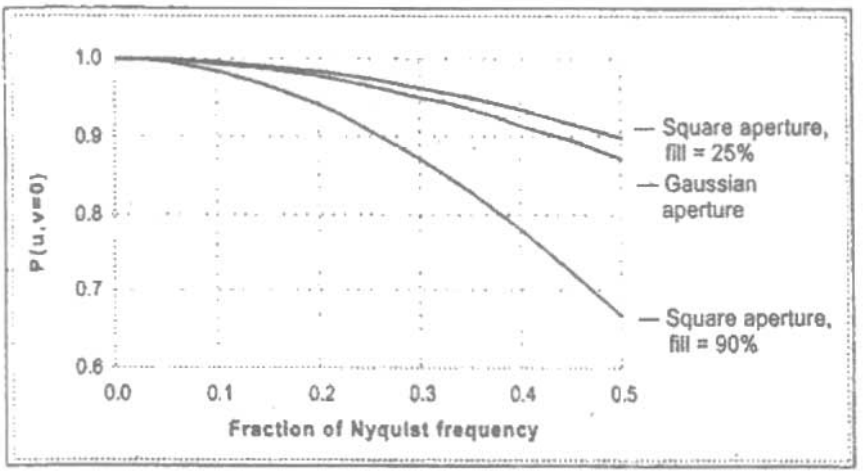
Trong đó *NA* là cửa chập số học của các thấu kính, λ là độ dài bước sóng ngắn nhất được sử dụng cùng với những thấu kính đó [16]. Nếu các thấu kính không đạt được tiêu chuẩn nào trong ba tiêu chuẩn đã giả định trước, nó vẫn sẽ có dải tần giới hạn nhưng có các tần số cắt thấp hơn so với những tần số đã cho trước trong phương trình (13). Khi làm việc với tần số *F* quang học thay vì với *NA* trong không khí (nơi có chỉ số khúc xạ là 1.0), phương trình (13) trở thành:

$$u_c = v_c = \frac{2}{\lambda}\left(\frac{1}{\sqrt{4F^2 + 1}}\right) \qquad (14)$$

3.1.1. Cửa chập lấy mẫu

Cửa chập $p(x,y)$ được mô tả ở trên chỉ gây ra một hiệu ứng nhỏ trên tín hiệu sau cùng nếu các điều kiện trong hai phương trình (12) và (13) được thỏa mãn. Giả sử như khoảng cách giữa các mẫu X_0 bằng Y_0 và cửa chập lấy mẫu không rộng hơn X_0, thì hiệu ứng xảy ra trên toàn phổ (do ứng xử của đại lượng $A(u,v)P(u,v)$ gây ra theo phương trình (9)) sẽ có dạng như trong hình 11 cho các cửa chập Gauss và cửa chập vuông.

Trong hình 11, phổ tín hiệu được ước lượng dọc theo một trục của phép biến đổi Fouier 2–chiều. Cửa chập Gauss trong hình này có bề rộng sao cho khoảng lấy mẫu X_0 chứa $\pm 3\sigma$ (99.7%) của phân số Gauss. Các cửa chập chữ nhật có bề rộng thỏa một trong số chúng chiếm 95% khoảng lấy mẫu, trong khi các cửa chập còn lại chiếm 50% khoảng lấy mẫu. Độ rộng 95% này sẽ được chuyển đổi thành một *hệ số làm đầy* 90%, còn độ rộng 50% thành *hệ số làm đầy* 25%. *Hệ số làm đầy* được giới thiệu trong mục 5.5.2.



Hình 11. *Phổ cửa chập P(u, v = 0) cho các tần số bằng ½ tần số Nyquist*

3.2. ĐỘ TRÙ MẬT LẤY MẪU TRONG PHÂN TÍCH ẢNH

Các nguyên tắc chọn độ trù mật lấy mẫu trong phân tích ảnh có nhiều điểm khác biệt so với trong xử lý ảnh. Điểm khác biệt cơ bản nhất ở đây là sự số hóa các đối tượng trong ảnh thành một tập các pixel sẽ dẫn đến một dạng nhiễu lượng hóa không gian có dải tần không giới hạn. Sự xuất hiện của loại nhiễu này dẫn đến những kết quả sau đây trong việc chọn lựa độ trù mật lấy mẫu nhằm phục vụ cho việc đo diện tích và tính độ dài chu vi của một vùng hay đối tượng.

3.2.1. Lấy mẫu để đo diện tích

Chúng ta giả sử rằng cách lấy mẫu được tiến hành theo kiểu lấy mẫu vuông, tức là $X_0 = Y_0$, và thuật toán ước lượng diện tích là không thiên vị theo kiểu đếm số pixel. Khi đó giá trị CV (xem phương trình (70)) của phép đo diện tích có liên hệ như sau với độ trù mật lấy mẫu [17]"

$$2D: \lim_{S\to\infty} CV(S) = k_2 S^{-3/2} \quad 3D: \lim_{x\to\infty} CV(S) = k_3 S^{-2} \qquad (15)$$

Còn trong không gian D–chiều:

$$\lim_{S\to\infty} CV(S) = k_D S^{-(D+1)/2} \qquad (16)$$

trong đó S là số mẫu trên mỗi đường kính đối tượng. Trong không gian 2–chiều giá trị đo được là diện tích, trong không gian 3–chiều giá trị đó là thể tích, và trong không gian D–chiều nó là một siêu thể tích.

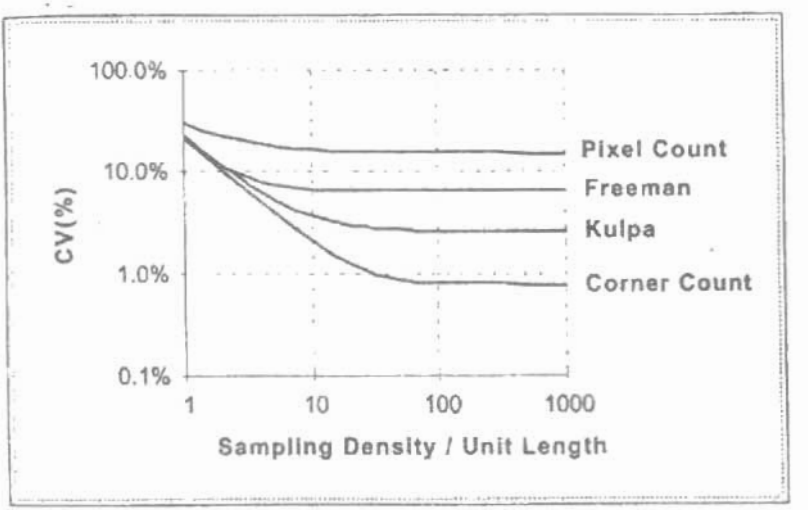
3.2.2. Lấy mẫu để đo độ dài

Một lần nữa chúng ta giả sử rằng cách lấy mẫu được tiến hành theo kiểu lấy mẫu vuông, tức là $X_0 = Y_0$, và các thuật toán ước lượng độ dài được dựa trên biểu diễn mã xích của Freeman (xem mục 6.6.1). Khi đó giá trị CV của phép đo độ dài sẽ có liên hệ với độ trù mật lấy mẫu (tính theo đơn vị độ dài) theo như trong hình 12 (xem [18] [19]).

Các đường cong trong hình 12 được phát triển trong ngữ cảnh là những đường thẳng, tuy nhiên chúng ta cũng có những kết quả tương tự cho đường cong và các đường biên đóng. Công thức ước lượng độ dài trong trường hợp này sử dụng biểu diễn mã xích của một đường thẳng và có sự tham gia của ba tham số theo như phương trình dưới đây.

$$L = \alpha \bullet N_e + \beta \bullet N_0 + \gamma \bullet N_c \qquad (17)$$

với N_e là các mã xích chẵn, N_0 là số các mã xích lẻ, còn N_c là số các góc. Các công thức cụ thể được cho trong bảng 5.



Hình 12. *Giá trị CV của phép đo độ dài ứng với nhiều thuật toán khác nhau*

3.2.3. Kết luận về lấy mẫu

Trong trường hợp ứng dụng cần xây dựng chủ yếu có liên quan đến xử lý ảnh, chúng ta nên chọn một độ trù mật lấy mẫu tuân theo lý thuyết tín hiệu cổ điển dựa trên định lý lấy mẫu Nyquist. Trong trường hợp ứng dụng lại thiên về phân tích ảnh, chúng ta nên chọn độ trù mật lấy mẫu dựa trên các tiêu chí đo đạc không thiên vị (*bias*) và chính xác (CV). Trong trường hợp không thể xác định rõ là ứng dụng thiên về phía nào, có lẽ nên chọn độ trù mật có giá trị cao hơn trong số hai giá trị độ trù mật có thể có được.

Bảng 5. *Các công thức ước lượng độ dài dựa trên các tham số có liên quan đến mã xích (N_e, N_o, N_e)*

Các hệ số	α	β	γ	
Công thức				**Tài liệu tham khảo**
Đếm số pixel	1	1	0	[18]
Freeman	1	$\sqrt{2}$	0	[11]
Kuipa	0.9481	$0.9481 * \sqrt{2}$	0	[20]
Đếm số góc	0.980	1.406	−0.091	[21]

Chương 4

CÁC LOẠI NHIỄU

Ảnh số thu nhận được từ các bộ cảm biến hiện đại có thể bị suy giảm chất lượng do rất nhiều nguồn nhiễu khác nhau. Với khái niệm nhiễu chúng ta muốn để cập đến những thay đổi ngẫu nhiên, thay vì những tác nhân làm lệch ảnh có tính tất định như hiệu ứng tạo bóng hay không đủ tiêu cự.

Trong chương này chúng ta sẽ khảo sát về nhiễu trên các ảnh được tạo nên từ các nguồn sáng trong những thiết bị quang điện tử hiện đại. Một cách cụ thể, thiết bị thu nhận ảnh là các máy chụp tích điện kép tiên tiến CCD (Charge–Coupled Device) với các photon sản sinh ra những điện tử thường được gọi là photoelectron. Mặc dù vậy, hầu hết các quan sát được trình bày ở đây về nhiễu, cũng như các nguyên nhân khác nhau của chúng đều có thể áp dụng được cho các mô hình lấy ảnh khác.

4.1. NHIỄU PHOTON

Khi tín hiệu được quan sát là ánh sáng, bản chất lượng tử của ánh sáng giữ vai trò hết sức quan trọng trong việc hình thành nhiễu. Một photon có bước sóng $\lambda = 500$nm mang năng lượng $E = h\nu = hc/\lambda = 3.97\times10^{-19}$ *jun*. Với các photon như vậy, các máy chụp CCD hiện đại có đủ nhạy của máy chụp trong mục 5.2). Bài toán nhiễu xuất hiện ở đây do bản chất thống kê của quá trình sản sinh ra các photon. Chúng ta không thể giả định được rằng ứng với mỗi pixel cho trước,

trong hai thời khoảng quan sát liên tiếp nhưng độc lập có độ dài T, lượng photon đếm được là như nhau. Do quá trình tạo photon được điều khiển bởi các qui luật trong vật lý lượng tử, cách tốt nhất để khảo sát nhiễu là quan tâm đến lượng photon trung bình đếm được trong một cửa sổ quan sát cho trước. Phân phối xác suất cho – photon trong một cửa sổ quan sát có độ dài T giây được biết là có dạng phân phối Poisson:

$$P(p|\rho, T) = \frac{(\rho T)^p e^{-\rho T}}{p!} \qquad (18)$$

trong đó ρ là tham số tốc độ/ cường độ được đo bằng đơn vị số photon/giây. Chúng ta cần lưu ý rằng ngay cả khi không có nguồn nhiễu nào khác trong chuỗi ảnh, thì do các dao động thống kê trong việc đếm số lượng photon trên một khoảng hữu hạn T vẫn có một tỷ số tín hiệu – nhiễu hữu hạn SNR tồn tại. Bằng cách sử dụng công thức tính SNR thích hợp theo phương trình (29), và để ý là giá trị trung bình cùng với độ lệch chuẩn trong phân phối Poisson liên hệ với nhau theo công thức:

Giá trị trung bình = ρT

Tiến trình Poisson·

$$\sigma = \sqrt{\rho T} \qquad (19)$$

Chúng ta sẽ có giá trị SNR được xác định bởi:

$$\text{Nhiễu photon – SNR} = 10 \log_{10}(\rho T)\text{dB} \qquad (20)$$

Với nhiễu photon, ba giả định truyền thống về mối quan hệ giữa tín hiệu và nhiễu sẽ không còn đúng nữa. Một cách cụ thể là:

– Nhiễu photon không độc lập với tín hiệu;

– Nhiễu photon không có dạng Gauss;

– Nhiễu photon không cộng tính.

Với những tín hiệu cực sáng giá trị – vượt quá 10^5, các dao động nhiễu do tính thống kê photon gây ra có thể bỏ qua được nếu bộ cảm biến ảnh có mức bão hòa đủ cao. Điều này sẽ được bàn đến kỹ lưỡng hơn trong mục 5.3 và cụ thể là trong phương trình (29).

4.2. NHIỄU NHIỆT

Chúng ta cần để ý rằng nguồn điện tử ngẫu nhiên trong một giếng CCD là nguồn năng lượng nhiệt. Các điện tử có thể tự giải phóng chúng ta khỏi vật liệu CCD nhờ dao động nhiệt và sau đó bị bắt lại trong giếng CCD. Những điện tử như vậy rất khó phân biệt khi so với các photoeletron thật sự. Bằng cách làm mát chip CCD, chúng ta có thể giảm được đáng kể số lượng các "điện tử nhiệt" vốn là nguyên nhân dẫn đến nhiễu nhiệt hay hiện tượng dòng tối. Khi thời gian tích hợp T tăng lên, số lượng các điện tử nhiệt cũng sẽ tăng theo. Phân phối xác suất của chúng cũng là một tiến trình Poisson có tham số tốc độ là hàm tăng theo nhiệt độ. Có nhiều kỹ thuật làm mát khác nhau để loại bỏ bớt dòng tối, trong số đó thường bao gồm việc ước lượng *dòng tối trung bình* ứng với một khoảng thời gian tích hợp cho trước, sau đó loại thời gian này ra khỏi các giá trị pixel của CCD trước khi tiến hành chuyển đổi tương tự – số. Mặc dù phương pháp này giúp làm giảm giá trị trung bình của dòng tối, nó lại không thể giúp làm giảm độ lệch chuẩn của dòng tối, hơn thế nữa nó còn làm giảm thang độ sáng động của tín hiệu.

4.3. NHIỄU ĐIỆN TỪ TRÊN CHIP

Nhiễu này có nguồn gốc trong quá trình đọc tín hiệu từ bộ cảm biến thông qua FET (Fielf Effect Transistor) của một chip CCD. Dạng tổng quát của độ trù mật phổ năng lượng ứng với nhiễu đọc ra là:

$$\textit{Nhiễu đọc ra} - S_{nn}(\omega) \propto \begin{cases} \omega^{-\beta} & \omega > \omega_{mm} \quad \beta > 0 \\ k & \omega_{min} < \omega < \omega_{max} \\ \omega^{\alpha} & \omega > \omega_{max} \quad \beta > 0 \end{cases} \qquad (21)$$

trong đó α và β là các hằng số, ω là tần số bán kính chuyển tín hiệu từ chip CCD ra thế giới bên ngoài. Ở các mức đọc ra cực thấp ($\omega < \omega_{min}$) nhiễu sẽ có tính chất *1/f*. Loại nhiễu này có thể được làm giảm đến mức quản lý được nhờ sự chọn lựa các tốc độ đọc ra thích hợp nhất. Tuy nhiên ở mức tín hiệu cực thấp (xem phương trình (20)), nhiễu đọc ra có thể vẫn trở thành một thành phần đáng kể trong giá trị SNR toàn thể [22].

4.4. NHIỄU KTC

Nhiễu liên kết với điện trở cổng của một FET được gọi là nhiễu KTC. Đây là nhiễu hoàn toàn không thể bỏ qua được và có giá trị đầu ra RMS về điện thế được cho bởi công thức sau:

$$\textit{Nhiễu KTC (điện thế)} - \sigma_{KTC} = \sqrt{\frac{kT}{C}} \qquad (22)$$

trong đó C là trở kháng chuyển cổng FET, k là hằng số Boltzmann, còn T là nhiệt độ tuyệt đối của chip CCD (được đo theo độ *K*). Bằng cách sử dụng các quan hệ $Q = C \bullet V = N_e \bullet e$

giá trị đầu ra RMS của nhiễu KTC tính theo số các photoelectron (N_{e^-}) sẽ được xác định bởi:

$$\textit{Nhiễu KTC (điện tử)} - \sigma_{N_{e^-}} = \frac{\sqrt{kTC}}{e^-} \qquad (23)$$

trong đó e^- là điện tích nguyên tố. Với C = 0.5pF và T = 233K chúng ta có N_e = 252 điện tử. Giá trị này là nhiễu "một lần" trên mỗi pixel xuất hiện trong khoảng thời gian đọc tín hiệu ra, do vậy nó độc lập với khoảng thời gian tích hợp (xem mục 4.1 và 5.7). Các thiết kế điện tử thích hợp có sử dụng kỹ thuật lấy mẫu kép tương quan có thể triệt tiêu hoàn toàn nhiễu KTC này [22].

4.5. NHIỄU KHUẾCH ĐẠI

Nhiễu loại này thường có tính chất cộng tính, có phân bô Gauss, và độc lập có tín hiệu. Trong những thiết bị điện tử được thiết kế tốt và hiện đại, nhiễu khuếch đại nhìn chung có thể bỏ qua được. Ngoại lệ phổ biến nhất trong số chúng là các máy chụp hình màu, với việc sử dụng khuếch đại trong kênh màu lục nhiều hơn ở kênh màu lam hay đỏ, từ đó dẫn đến có nhiều nhiễu hơn trong kênh màu lục (xem thêm mục 5.6).

4.6. NHIỄU LƯỢNG HÓA

Nhiễu lượng hóa xuất hiện trong quá trình lượng hóa độ lớn tín hiệu và ở công đoạn chuyển đổi tương tự – số thông qua các bộ ADC. Nó có tính chất cộng tính và độc lập với tín hiệu khi số các mức lượng hóa $L \geq 16$. Điều này tương đương với B – 4 bit (xem mục 2.1). Với một tín hiệu đã được chuyển đổi thành dạng điện tử và do vậy có các giá trị lớn nhất cũng như nhỏ nhất, phương trình (72) là công thức phù hợp nhất

cho việc xác định giá trị SNR. Nếu bộ chuyển đổi ADC được điều chỉnh sao cho 0 tương ứng với giá trị điện tử nhỏ nhất, còn $2^B - 1$ tương ứng với giá trị điện tử lớn nhất thì:

$$\textit{Nhiễu lượng hóa} - SNR = 6B + 11dB \qquad (24)$$

Với B ≥ 8bit, điều này có nghĩa là $SNR \geq 59d$. Nhiễu loại này có thể bỏ qua được khi giá trị SNR tổng thể của một hệ thống hoàn toàn bị thống trị bởi một loại nhiễu cho SNR *nhỏ nhất*. Trong các máy chụp CCD đó chính là nhiễu photon.

Chương 5

CÁC THIẾT BỊ THU ẢNH VÀ HIỂN THỊ ẢNH

Các thiết bị thu nhận ảnh ngày nay bao gồm rất nhiều chủng loại và được chế tạo dựa trên rất nhiều công nghệ khác nhau, từ đơn giản đến phức tạp tùy theo mục đích sử dụng. Chúng ta không có tham vọng trình bày hết về các thiết bị thu ảnh trong chương này, mà chỉ giới hạn trong việc khảo sát các tính năng kỹ thuật có liên quan đến loại máy chụp phổ biến nhất được dùng trong xử lý ảnh là máy chụp CCD.

5.1. TÍNH TUYẾN TÍNH

Mối quan hệ thường được mong muốn giữa một tín hiệu vật lý đầu vào (các photon) với một tín hiệu đầu ra (điện thế) là mối quan hệ tuyến tính. Một cách hình thức điều này có nghĩa là (như trong phương trình (80) nếu chúng ta có hai ảnh a và b, cùng với hai hằng số phức tùy ý w_1 và w_2, thêm vào đó là một đáp ứng camera tuyến tính thì:

$$c = \Re\{w_1 a + w_2 b\} = w_1\Re\{a\} + w_2\Re\{b\} \qquad (25)$$

trong đó $\Re\{\bullet\}$ là đáp ứng của camera, c là kết xuất của camera đó. Trong thực tế mối quan hệ giữa ảnh đầu vào a và ảnh đầu ra c thường được xác định bởi:

$$c = gain \bullet a^{\gamma} + offset \qquad (26)$$

Ở đây γ là giá trị gamma của thiết bị ghi ảnh. Với một hệ ghi ảnh tuyến tính thật sự chúng ta phải có $\gamma = 1$ và $offset = 0$, nhưng do giá trị offset trong thực tế gần như không bao giờ

bằng 0, chúng ta phải chú ý bù trừ lại yếu tố này nếu như định rút trích ra các số đo về mặt cường độ sáng. Các kỹ thuật bù trừ được khảo sát trong mục 7.1.

Bảng 6 sau đây liệt kê một số giá trị tiêu biểu của γ. Các camera hiện đại thường có khả năng chuyển đổi qua lại giữa các giá trị γ.

Bảng 6: *So sánh các giá trị γ*

Bộ cảm biến	Bề mặt	γ	Lợi điểm khi sử dụng
Chip CCD	Cilicon	1.0	Tính tuyến tính
Ống vidicon	Sb_2S_3	0.6	Nén thang độ sáng động – Các cảnh có độ tương phản cao
Phim	Silver halide	<1.0	Nén thang độ sáng động – Các cảnh có độ tương phản cao
Phim	Silver halide	> 1.0	Mở rộng thang độ sáng động – Các cảnh có độ tương phản thấp

5.2. ĐỘ NHẠY

Có hai cách để mô tả độ nhạy của một camera. Cách thứ nhất là đi xác định lượng nhỏ nhất các photoelectron có thể dò được qua đó thu được độ nhạy *tuyệt đối*. Cách thứ hai là mô tả số các photoelectron cần thiết để chuyển từ một mức sáng sang một mức sáng kế tiếp, hay còn được gọi là thay đổi một *đơn vị tương tự – số* (ADU). Cách làm này dẫn đến độ nhạy tương đối của camera.

5.2.1. Độ nhạy tuyệt đối

Để xác định độ nhạy tuyệt đối chúng ta cần đặc trưng hóa một camera dưới dạng nhiễu của nó. Nếu nhiễu tổng cộng

có chẳng hạn – 100 photoelectron, thì để bảo đảm tính dò được của một tín hiệu chúng ta có thể nói rằng ở mức 3σ, tín hiệu dò được tối thiểu (hay độ nhạy tuyệt đối) sẽ là 300 photoelectron. Nếu mọi nguồn nhiễu đã liệt kê trong chương 4, ngoại trừ nhiễu photon, đều có thể giảm đến mức bỏ qua được, chúng ta sẽ có được độ nhạy tuyệt đối không vượt quá 10 photoelectron với các công nghệ chế tạo CCD hiện đại.

5.2.2. Độ nhạy tương đối

Với định nghĩa đã nêu ở trên về độ nhạy tương đối *S* trong trường hợp tuyến tính (phương trình (26)) và $\gamma = 1$ chúng ta sẽ có:

$$S = \frac{1}{\text{gain}} = \text{gain}^{-1} \qquad (27)$$

Quá trình đo *độ nhạy* hay *độ lợi (gain)* có thể thực hiện theo hai cách khác nhau sau đây.

Cách 1

Theo phương trình (26), nếu tín hiệu đầu vào *a* có thể được điều khiển chính xác thông qua thời gian tích hợp hoặc cường độ (nhờ các bộ lọc trù mật trung tính), khi đó độ lợi có thể tính được qua hệ số góc của đường thẳng kết quả. Để chuyển đổi giá trị này sang các đơn vị đo mong muốn, chúng ta cần phải sử dụng một nguồn chuẩn phát ra một lượng photon biết trước lên bộ cảm biến camera, đồng thời độ hiệu quả lượng tử (η) cũng phải được cho trước. Ở đây độ hiệu quả lượng tử cho biết số photoelectron được sản sinh ra trên mỗi photon ứng với một độ dài sóng cho trước. Một cách tổng quát chúng ta có $0 \le \eta(\lambda) \le 1$.

Cách 2

Trong trường hợp hiệu ứng giới hạn của camera chỉ là nhiễu photon (xem mục 6.1), chúng ta còn có một kỹ thuật khác dễ cài đặt hơn để xác định được độ nhạy của camera. Bằng cách sử dụng phương trình (19), (26), (27) và sau khi bù trừ độ dịch chuyển offset (xem mục 7.1), độ nhạy đo được từ một cảnh c sẽ là:

$$S = \frac{E\{c\}}{Var\{c\}} = \frac{m_c}{s_c^2} \qquad (28)$$

trong đó m_c và s_c được xác định thông qua các phương trình (66) và (68)

Bảng liệt kê các dữ liệu đo được ứng với năm kiểu camera CCD sản xuất năm 1995. Trong bảng này, với một camera CCD kiểu khoa học (C–1), chúng ta chỉ cần có 8 photoelectron (gần 16 photon) dùng để tách biệt hai mức xám trong biểu diễn số hóa của ảnh. Với một video camera ít tốn kém hơn (C–5), chúng ta cần phải có khoảng 256 photoelectron (gần 512 photon) để tách biệt hai mức xám. Như vậy độ nhạy của các camera biến đổi rất rõ rệt khi đi từ những sản phẩm thông thường cho đến chuyên dụng.

Bảng 7. *Các giá trị độ nhạy ứng với nhiều loại camera CCD. Chúng ta cần chú ý rằng có nhiều hơn một camera có giá trị thấp hơn S*

Camera Nhãn	Pixel	Cỡ pixel µm x µm	Nhiệt độ K	S e^-/ADU	Số bit
C–1	1320 x 1035	6.8 x 6.8	231	7.9	12
C–2	578 x 385	22.0 x 22.0	227	9.7	16

C-3	1320 x 1035	6.8 x 6.8	293	48.1	10
C-4	576 x 384	23.0 x 23.0	238	90.9	12
C-5	756 x 581	11.0 x 5.5	300	109.2	8

5.3. SNR

Như đã đề cập đến trong chương 4, trong các hệ thống camera hiện đại nhiễu thường bị giới hạn bởi:

– Nhiễu khuếch đại trong trường hợp các camera màu.

– Nhiễu nhiệt. Nhiễu này tự giới hạn bởi nhiệt độ chip K và thời gian lộ sáng T, hoặc

– Nhiễu photon. Nhiễu này bị giới hạn bởi tốc độ sản sinh photon – và thời gian lộ sáng T.

5.3.1 Nhiễu nhiệt (dòng tối)

Bằng cách sử dụng những kỹ thuật làm mát dựa trên các phần tử làm mát Peltier, nhiệt độ chip có thể giảm xuống còn ở mức từ 230 đến 250 độ K, và khi đó tốc độ sản sinh ra các điện tử nhiệt sẽ giảm xuống. Để đo nhiễu nhiệt, chúng ta có thể chú ý đến khoảng thời gian cần thiết (tính theo giây) để tạo ra một lượng đủ lớn các điện tử nhiệt khi đi từ một mức sáng sang một mức sáng kế tiếp (một ADU) với sự thiếu vắng các photoelectron. Điều kiện cuối cùng này (sự thiếu vắng các photoelectron) chính là nguyên nhân dẫn đến tên gọi dòng tối. Các đặc trưng của nhiễu nhiệt ứng với năm loại camera tiêu biểu được cho trong bảng 8 dưới đây.

Bảng 8. *Các đặc trưng của nhiễu nhiệt*

Camera Nhãn	Nhiệt độ K	Dòng tối giây /ADU
C–1	231	526.3
C–2	227	0.2
C–3	293	8.3
C–4	238	2.4
C–5	300	23.3

Trong bảng trên, video camera (C–5) có sự giảm thiểu dòng tối trên chip (xem mục 4.2). Khi hoạt động ở nhiệt độ phòng camera này cần không nhiều hơn 20 giây để tạo ra một thay đổi ADU do nhiễu nhiệt. Điều này có nghĩa là với tốc độ phát hình từ 25 đến 30 ảnh/ giây (xem bảng 3) của camera, nhiễu nhiệt có thể bỏ qua được.

5.3.2. Nhiễu photon

Từ phương trình (20) chúng ta nhận thấy rằng nên tăng giá trị SNR bằng cách răng thời gian tích hợp ảnh để "bắt" được nhiều photon hơn. Các pixel trong một camera CCD có dung lượng giếng hữu hạn C chính là giá trị SNR tối đa/pixel:

Nhiễu photon có dung lượng giới hạn –

$$\text{SNR} = 10 \log_{10}(C) \text{ dB} \qquad (29)$$

Bảng 9 cho chúng ta các dữ liệu có liên quan đến nhiễu photon đo được trong thực tế cũng như dựa trên lý thuyết. Chúng được thực hiện trên năm loại camera tiêu biểu.

Bảng 9. *Các đặc trưng của nhiễu photon*

Camera Nhãn	C Số e	Giá trị SNR lý thuyết (dB)	Giá trị SNR đo được (dB)	Cỡ pixel µm x µm	Độ sâu giếng Số $e^-/\mu m^2$
C-1	32000	45	45	6.8 x 6.8	692
C-2	340000	55	55	22.0 x 22.0	702
C-3	32000	45	43	6.8 x 6.8	692
C-4	400000	56	52	23.0 x 23.0	756
C-5	40000	46	43	11.0 x 5.5	661

5.4. TẠO BÓNG

Trong tất cả các hệ thống lấy ảnh chúng ta đều mong muốn thu được kết quả là với một ảnh vật lý đầu vào $a(x, y)$ không đổi thì phiên bản số hóa của ảnh cũng không đổi. Tuy nhiên trong thực tế khó lòng đạt được kết quả đó, do hầu hết các hệ lấy ảnh đều tạo ra bóng, với nguồn bóng có thể nằm ngoài camera do độ chiếu sáng trong cảnh lấy ảnh, hoặc do chính bản thân camera tạo nên khi giá trị *gain* và *offset* thay đổi từ pixel này sang pixel khác. Mô hình tạo bóng được xác định thông qua phương trình sau:

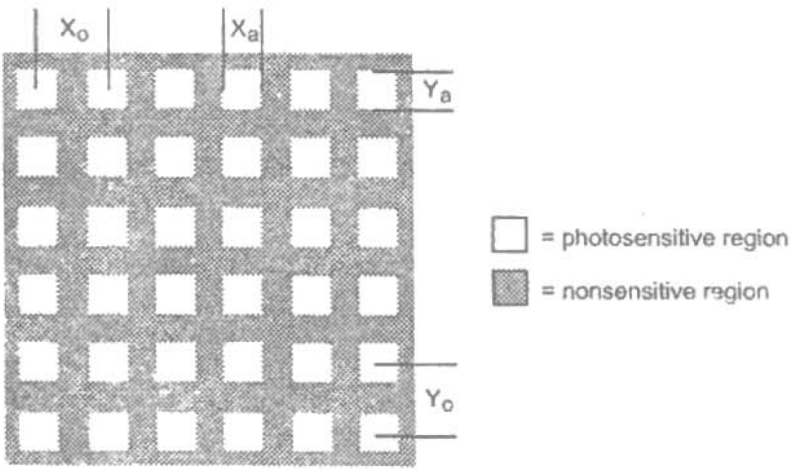
$$c[m, n] = gain[m, n] \bullet a[m, n] + offset[m, n] \quad (30)$$

Trong phương trình trên $a[m, n]$ là ảnh số hóa sẽ được ghi nhận nếu như không có hiện tượng tạo bóng trong ảnh, tức là $a[m, n] = constant$. Các kỹ thuật có liên quan đến việc làm giảm hoặc khử bóng có thể tìm thấy trong mục 7.1

5.5. DẠNG PIXEL

Mặc dù các pixel đã được định nghĩa trong hình 2 là luôn có dạng hình vuông và phủ toàn bộ ảnh liên tục, các tính

chất hình học của một hệ thống số hóa/ camera vẫn có những ảnh hưởng nhất định lên trên chúng. Hình 13 cho chúng ta định nghĩa của nhiều tham số có liên quan đến camera/ bộ số hóa và có ảnh hưởng đối với việc hình thành các pixel ảnh.



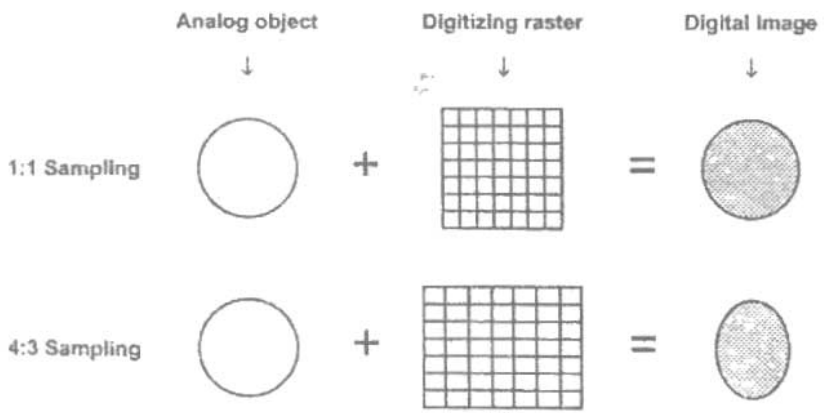
Hình 13. *Các tham số dạng pixel*

Trong hình 13, các tham số X_0 và Y_0 là khoảng cách giữa các tâm pixel. Chúng biểu thị các khoảng cách lấy mẫu theo phương trình (8). Các tham số X_a và Y_a là các kích thước của phần bề mặt camera nhạy với ánh sáng. Như đã đề cập trong mục 2.3, các bộ số hóa video khác nhau có thể có nhiều giá trị X_0 trong khi vẫn có chung một giá trị Y_0.

5.5.1. Các pixel vuông

Như đã đề cập đến trong chương 3, việc lấy mẫu vuông sẽ dẫn đến $X_0 = Y_0$ hay $X_0/Y_0 = 1$. Trong thực tế các máy thu ảnh có $X_0/Y_0 = 1.1$ hoặc $X_0/Y_0 = 4/3$ cũng rất phổ biến, đặc

biệt là tỷ số 4/3 cho truyền hình thương mại. Độ rủi ro ứng với các pixel không vuông nằm ở chỗ khi các đối tượng đẳng hướng được quét vào thì trên màn hình camera chúng vẫn xuất hiện dưới dạng đẳng hướng, còn khi phân tích dựa trên các yếu tố có liên quan đến chiều dài và chiều rộng chúng sẽ cho ra các kết quả không đẳng hướng. Điều này được minh họa trên hình 14.



Hình 14. *Tác động của các pixel không vuông*

5.5.2. Hệ số làm đầy

Trong các camera CCD hiện đại có thể có tình huống một phần bề mặt của camera không nhạy sáng và được sử dụng để ngăn ngừa hiện tượng *blooming*. Hiện tượng blooming xuất hiện khi một giếng CCD được làm đầy (xem bảng 9) và các photoelectron phụ thêm tràn vào các giếng CCD kế cận. Các vùng anti–blooming giữa các site CCD thường được sử dụng để ngăn ngừa hiện tượng này xảy ra với giả định là sẽ có một phần nhỏ các photon đến bị mất đi khi

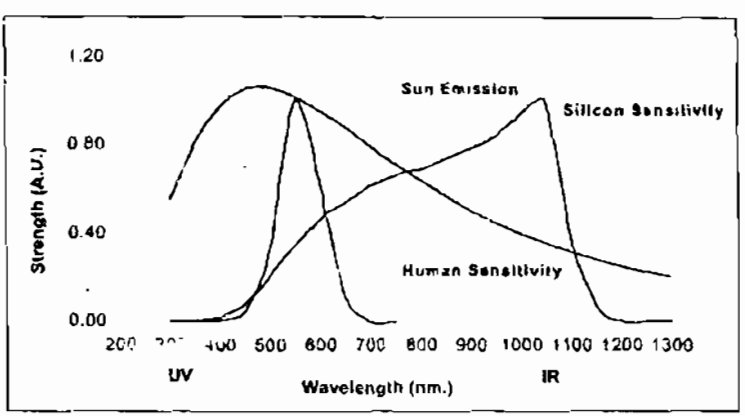
đập vào phần không nhạy sáng của chip CCD. Phần rất nhỏ bề mặt nhạy sáng được gọi là *hệ số làm đầy (fill–factor)* và được xác định theo công thức sau:

$$\text{fill} - \text{factor} = \frac{X_a \bullet Y_a}{X_0 \bullet Y_0} \times 100\% \qquad (31)$$

Với giá trị *fill–factor* càng lớn, chip càng có khả năng bắt được nhiều ánh sáng cho đến mức tối đa là 100%, do vậy có thể giúp cải thiện giá trị SNR. Tuy nhiên giá trị của hệ số làm đầy càng lớn cũng dẫn đến độ trơn không gian nhiều hơn do hiệu ứng cửa chập đã mô tả trong mục 3.1.1. Điều này được minh họa trong hình 17.

5.6. ĐỘ NHẠY PHỔ

Các bộ cảm biến được sử dụng trong camera và phim ảnh thường không có độ nhạy đồng nhất với tất cả các bước sóng, ánh sáng. Độ nhạy phổ cho một bộ cảm biến CCD được thể hiện trong hình 15.



Hình 15: *Các đặc trưng phổ của silicon, mặt trời, và hệ thị giác của con người. UV viết tắt cho* ultraviolet *(tia cực tím), còn IR viết tắt cho* infra–red *(tia hồng ngoại).*

Độ nhạy cao của silicon trong miền hồng ngoại cho thấy rằng, với các ứng dụng trong đó camera CCD được sử dụng như nguồn cung cấp ảnh cho quá trình phân tích và xử lý, việc sử dụng bộ lọc đóng khối IR cần phải được xem xét cẩn thận. Bộ lọc này sẽ đóng khối các bước sóng trên 750 nm, do vậy ngăn cản hiện tượng "sương mù" trong ảnh có được từ các bước sóng dài hơn trong ánh sáng mặt trời. Nói cách khác, một camera CCD có thể dùng để chế tạo nên một bộ cảm biến tuyệt hảo cho các bước sóng ở gần vùng hồng ngoại, trong khoảng từ 750nm đến 100nm.

5.7. TỐC ĐỘ CHỚP (THỜI GIAN TÍCH HỢP)

Khoảng thời gian để ảnh được phơi ra ánh sáng – hay nói cách khác là khoảng thời gian thu thập các photon – có thể thay đổi dựa trên chủng loại camera hoặc dựa trên định dạng video được sử dụng (xem bảng 3). Nó thường được gọi bằng một trong hai thuật ngữ là tốc độ chớp hoặc thời gian tích hợp.

5.7.1. Tốc độ chớp trên các video camera

Trên các video camera CCD kiểu thương mại tốc độ chớp thường ở mức 500 ns. Người ta còn sử dụng các tốc độ qui ước phổ biến là 33.37 mn cho hệ NTSC và 40.0 ms cho các hệ PAL & SECAM. Các giá trị tốc độ chớp cao cỡ 30 giây cũng có thể đạt được với một số camera nhất định, tuy nhiên làm như vậy chúng ta đã hy sinh tính liên tục của dòng dữ liệu ảnh video để có được một ảnh tích hợp duy nhất thay vì nhiều ảnh.

5.7.2. Tốc độ chớp trên các camera khoa học

Với các camera khoa học, tốc độ chớp có thể đạt mức 500ns. Với công nghệ làm mát chip dựa trên kỹ thuật làm

mát Peltier hoặc làm mát bằng nitơ lỏng, chúng ta hoàn toàn có thể đạt được thời gian tích hợp vượt quá một giờ đồng hồ trên các thiết bị chuyên dụng loại này.

5.8. TỐC ĐỘ ĐỌC RA

Tốc độ đọc dữ liệu từ chip cảm biến được gọi là tốc độ đọc ra. Tốc độ đọc ra cho các video camera chuẩn phụ thuộc vào nhiều tham số của bộ bắt ảnh cũng như chính bản thân camera. Với định dạng video chuẩn (xem mục 2.3), tốc độ đọc ra sẽ được cho bởi:

$$R = \left(\frac{\text{số ảnh}}{\text{giây}}\right)\left(\frac{\text{số dòng}}{\text{ảnh}}\right)\left(\frac{\text{số điểm ảnh}}{\text{dòng}}\right) \qquad (32)$$

Mặc dù đơn vị mô tả thích hợp nhất cho tốc độ đọc ra là số pixel/ giây, người ta vẫn hay sử dụng thuật ngữ Hz để làm đơn vị biểu diễn. Với một camera sử dụng các pixel vuông (xem mục 5.5) chúng ta có bảng các tham số tương ứng sau.

Bảng 10. *Các tốc độ đọc ra ứng với các định dạng video chuẩn*

Định dạng	Số dòng/ giây	Số pixel/ dòng	R (MHz)
NTSC	15,750	(4/3)*525	~ 11.0
PAL/ SECAM	15,625	(4/3)*625	~ 13.0

Chúng ta cần chú ý rằng các số liệu trong bảng 10 chỉ có tính xấp xỉ. Giá trị chính xác cho một hệ thống sử dụng pixel vuông đòi hỏi tri thức chính xác về cách bộ số hóa video lấy mẫu mỗi dòng.

Các tốc độ đọc ra được sử dụng trong video camera hàm ý nhiễu điện tử đã mô tả trong mục 6.3 xuất hiện trong vùng

phổ nhiễu (phương trình (21)) mô tả bởi $\omega > \omega_{max}$, là nơi năng lượng nhiễu tăng theo tần số. Do vậy nhiễu đọc ra có thể trở thành đáng kể trong các video camera.

Các camera khoa học thường sử dụng tốc độ đọc ra thấp hơn để làm giảm nhiễu đọc ra. Các tốc độ đọc ra cụ thể thường được sử dụng là 20 kHz, 500 kHz, 1 MHz cho đến 8 MHz.

5.9. CÁC THIẾT BỊ HIỂN THỊ ẢNH

Các màn hình được sử dụng trong xử lý ảnh – đặc biệt là các màn hình dùng chung với máy tính – có một số tính chất giúp xác định chất lượng của ảnh sau cùng.

5.9.1. Tốc độ sinh động

Tốc độ sinh động được định nghĩa là số các ảnh hoàn chỉnh ghi được ra màn hình trong mỗi giây. Với các định dạng video chuẩn, tốc độ này là cố định theo các giá trị đã cho trong bảng 3, tức là 29.97 hay 25 ảnh/ giây. Với màn hình máy tính, tốc độ sinh động có thể biến đổi từ 67 ảnh/ giây đến 785 ảnh/ giây. Ở các giá trị trên mức 60 ảnh/ giây hiệu ứng rung thị giác có thể bỏ qua cho hầu hết các mức chiếu sáng.

5.9.2. Hệ số đan xen

Nhằm ngăn chặn hiện tượng rung thị giác ở các tốc độ làm tươi dưới 60 ảnh/ giây, màn hình hiển thị cần phải được đan xen như đã mô tả trong mục 2.3. Độ đan xen tiêu chuẩn cho các hệ thống video là 2:1. Do kỹ thuật đan xen là không cần thiết khi tốc độ làm tươi của màn hình vượt quá 60 ảnh/ giây, nên trong trường hợp đó hệ số đan xen sẽ là 1:1. Nói cách khác, các dòng sẽ được vẽ theo thứ tự thông thường từ dòng 1, 2, 3, 4, ..., cho đến dòng N.

5.9.3. Độ phân giải

Mặc dù được tạo ra từ các vùng có diện tích hữu hạn trong cảnh quan ban đầu (xem mục 5.1 và 5.5), khi được lưu trữ trong bộ nhớ máy tính các pixel vẫn có thể được xem là các điểm toán học không có kích thước vật lý. Tuy nhiên, khi được hiển thị, khoảng cách giữa các điểm này phải được lấp đầy. Quá trình này diễn ra được do sự làm việc của ống CRT với vùng chiếu sáng có kích thước hữu hạn. Phần vùng sáng tạo bởi CRT có phân phối xấp xỉ phân phối Gauss và số lượng của chúng phụ thuộc vào chất lượng của màn hình. Người ta vẫn thường đặc tả một hệ thống hiển thị theo kiểu có độ phân giải 72 vùng sáng/inch (28.3 vùng sáng/cm). Con số này tương ứng với các qui ước chuẩn về in ấn, tuy nhiên trong trường hợp đối tượng hiển thị không phải là máy in thì chúng ta còn có thể sử dụng độ phân giải cao hơn, ví dụ như 30 vùng sáng/ cm chẳng hạn.