ĐẠI HỌC CÔNG NGHỆ THÔNG TIN

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH

**KHOA KHOA HỌC MÁY TÍNH**



**BÁO CÁO ĐỒ ÁN**

**MÔN TÍNH TOÁN ĐA PHƯƠNG TIỆN**

**ĐỀ TÀI:**

**CÁC THUẬT TOÁN NÉN ẢNH**

**HUFFMAN, LZW, JPEG STANDARD**

**Lớp: CS232.J11.KHTN**

**GV Lý Thuyết: Ngô Đức Thành**

**GV HDTH: Đỗ Văn Tiến**

**Nhóm thực hiện:**

**Nguyễn Trọng Nhân 15520570**

**Trần Nguyên Khánh 15520363**

**Nguyễn Văn Minh 15520488**

**Huỳnh Ngọc Thiên Trang 15520917**

Thành phố Hồ Chí Minh, ngày 28 tháng 12 năm 2018



**LỜI CẢM ƠN**

Trong suốt thời gian làm đề tài này, nhóm đã nhận được rất nhiều sự quan tâm giúp đỡ của các thầy. Với lòng biết ơn sâu sắc, nhóm xin gởi lời cảm ơn chân thành đến thầy Ngô Đức Thành và thầy Đỗ Văn Tiến – người đã tận tình dạy bảo và hướng dẫn nhóm chúng em.

Vì thời gian và năng lực hạn chế nên trong quá trình làm báo cáo khó tránh khỏi sai sót về nội dung, lỗi trình bày mong các Thầy bỏ qua. Đồng thời nhóm cũng mong nhận được ý kiến chia sẻ, đóng góp của các Thầy. Cuối cùng nhóm chúng em xin kính chúc quý các Thầy dồi dào sức khỏe và thành công trong sự nghiệp cũng như cuộc sống.

Trân trọng kính chào

**NHẬN XÉT CỦA GIẢNG VIÊN**

Mục lục

# PHÂN CÔNG CÔNG VIỆC

* 1. Phân công công việc.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Thành viên | Nguyễn Trọng Nhân | Trần Nguyên Khánh | Nguyễn Văn Minh | Huỳnh Ngọc Thiên Trang |
| Nội dung | - Thiết kế giao diện, tổng hợp các thuật toán thành một chương trình hoàn chỉnh. Tìm hiểu các lưu file nén.  - Viết và chỉnh sửa báo cáo. | - Code thuật toán nén ảnh chuẩn JPEG.  - Viết báo cáo. | - Code thuật toán nén ảnh LZW.  - Viết báo cáo | - Code thuật toán nén ảnh Huffman.  - Viết báo cáo. |

* 1. Đánh giá.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Thành viên | Nguyễn Trọng Nhân | Trần Nguyên Khánh | Nguyễn Văn Minh | Huỳnh Ngọc Thiên Trang |
| Nhận xét chung | - Hoàn thành công việc được giao.  - Có ý thức tránh nhiệm với nhóm và đồ án. Nhiệt tình hỗ trợ các thành viên.  - Đi làm việc nhóm đúng giờ, nghiêm túc.  - Giao diện đơn giản, thân thiện người dùng, đảm bảo đủ chức năng. | - Hoàn thành công việc được giao.  - Có ý thức tránh nhiệm với nhóm và đồ án.  - Đi làm việc nhóm đúng giờ, nghiêm túc.  - Thuật toán trả về kết quả chính xác. Tuy nhiên thời gian chạy thuật toán còn chậm | - Hoàn thành công việc được giao.  - Có ý thức tránh nhiệm với nhóm và đồ án.  - Đi làm việc nhóm đúng giờ, nghiêm túc.  - Thuật toán trả về kết quả chính xác. Tuy nhiên thời gian chạy thuật toán còn chậm | - Hoàn thành công việc được giao.  - Có ý thức tránh nhiệm với nhóm và đồ án.  - Đi làm việc nhóm đúng giờ, nghiêm túc.  - Thuật toán trả về kết quả chính xác. Thời gian chạy thuật toán nhanh. |
| Mức độ hoàn thành công việc. | 100% | 100% | 100% | 100% |

1. SƠ LƯỢC ĐỀ TÀI
   1. Giới thiệu

Ảnh là một trong những loại dữ liệu được dùng phổ biến để lưu trữ và truyền tải thông tin.

* 1. Các công việc cần thực hiện
* Tìm hiểu về ý tưởng và thuật toán nén ảnh: Huffman, LZW, chuẩn JPEG.
* Tìm hiểu cách đọc file

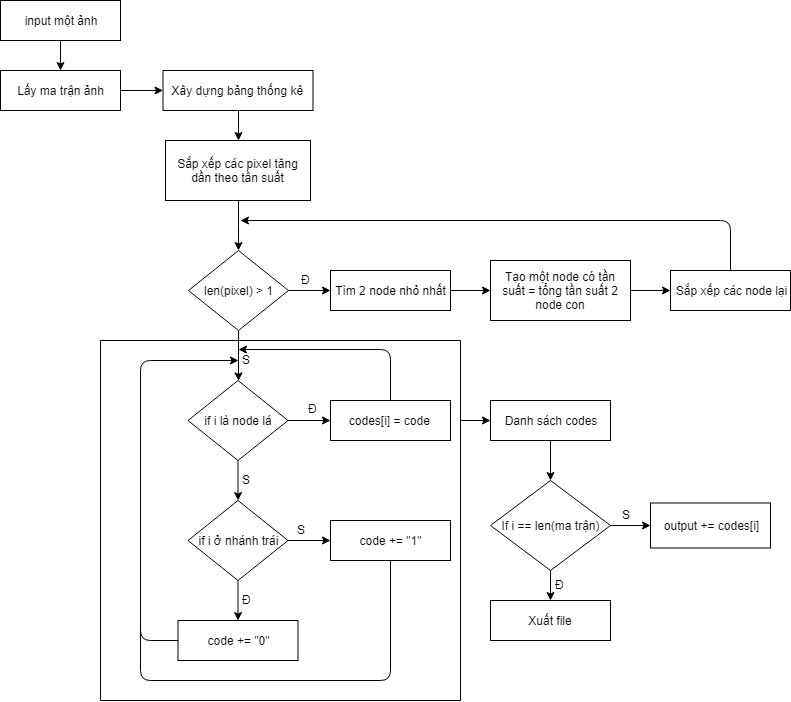
1. Thuật toán mã hóa Huffman
   1. Giới thiệu

* Trong khoa học máy tính và lý thuyết thông tin, mã hóa Huffman là một thuật toán mã hóa dùng để nén dữ liệu. Nó dựa trên bảng tần suất xuất hiện của các ký tự cần mã hóa để xây dựng một bộ mã nhị phân cho các ký tự đó sao cho dung lượng (số bit sau khi mã hóa là nhỏ nhất).
* Thuật toán được đề xuất bởi David A. Huffman khi ông còn là sinh viên Ph. D tại MIT, và công bố năm 1952 trong bài báo “A Method for the Construction of Minimum – Redundancy Codes”.
  1. Thuật toán
     1. Quá trình nén dữ liệu

Input: là một hình ảnh cần được nén

Output: Cây mã huffman

* Bước 1: Xây dựng bảng thống kế tần suất xuất hiện của các pixel
* Bước 2: Sắp xếp các pixel theo tần suất tăng dần
* Bước 3: Tìm 2 node pixel có tần suất xuất hiện thấp nhất và ghép lại thành 1 node với tần suất bằng tổng tần suất của 2 node pixel này
* Bước 4: Thêm node vừa tạo ở bước 3 vào danh sách có node.
* Bước 5: Gán bit 0 cho nhánh bên trái, nhánh bên phải gán bit 1.
* Bước 6: Tổng hợp tất cả các node lại với nhau để được tạo một cây mã nhị phân.

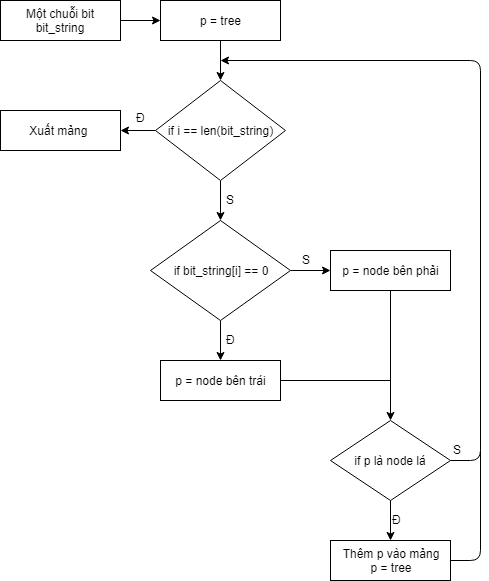


* + 1. Quá trình giải nén dữ liệu

Input: là một file binary của ảnh cầ được giải nén

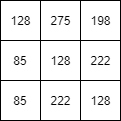
Output: là một ảnh

Truy vấn lại cây mã nhị phân vừa tạo ra ở quá trình nén dữ liệu, nếu bit là 0 thì truy hồi lại nhánh bên trái, nếu bit là 1 thì truy hồi lại nhánh bên phải. Quá trình được lặp lại cho đến node cuối.



* 1. Ví dụ

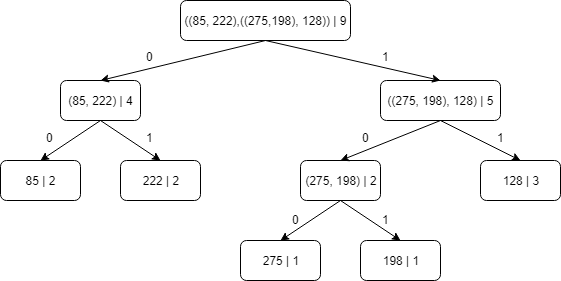
Đây là ma trận pixel 3x3:



Bảng tần suất xuất hiện của các pixel đã được sắp xếp:

|  |  |
| --- | --- |
| Pixel | Tần suất xuất hiện |
| 275 | 1 |
| 198 | 1 |
| 85 | 2 |
| 222 | 2 |
| 128 | 3 |

Cây mã huffman:



Bảng mã cây huffman:

|  |  |
| --- | --- |
| Pixel | Code |
| 275 | 100 |
| 198 | 101 |
| 85 | 00 |
| 222 | 01 |
| 128 | 11 |

* 1. Cài đặt

Các hàm chính của chương trình:

* get\_matrix\_image (): là hàm lấy ma trận pixel của hình ảnh.
* cal (): là hàm tính tần suất xuất hiện của các pixel.
* sortFreq (): hàm sắp xếp các pixel theo tần suất tăng dần.
* buildTree (): hàm ghép 2 node có tần suất xuất hiện thấp nhất và tạo một node cha của 2 node này thêm vào mảng node (node có tần suất xuất hiện và giá trị pixel).
* Tree (): hàm lấy giá trị pixcel của mảng node (loại bỏ tần suất xuất hiện của các node) tạo thành cây hoàn chỉnh.
* assignCodes (): hàm tìm mã code của từng pixel.
* encode () và read (): hàm thực hiện lưu cây mã huffman dưới dạng file binary.
* decode (): giải nén cây mã huffman.

1. d
2. d
3. d
4. d
5. d
6. d
7. Xây dựng giao diện
   1. s
8. s
9. s