

# Báo cáo bài tập giữa kì xử lý ảnh

Họ tên: Nguyễn Quốc Khánh – MSSV: 18020710

References: [Ciratefi: An RST-Invariant Template Matching](#)

## 1. Overview

Template matching là một kỹ thuật xử lý ảnh số (digital image) được sử dụng phổ biến để nhận dạng vật thể trong ảnh thông qua ảnh template (template). Một kỹ thuật xử lý ảnh hiệu quả phải có khả năng nhận diện được các vật thể đã trải qua một số phép biến đổi hình học. Círatefi (Circular, Radial and Template-Matching Filter) là một thuật toán template matching cho ảnh xám, có cơ chế xử lý các phép xoay, phóng to thu nhỏ, và dịch chuyển và khá mạnh đối với ảnh bị mờ (blur) và bị nén JPEG (JPEG-compression).

## 2. Kỹ thuật Círatefi

Círatefi là một thuật toán template matching với ảnh xám bao gồm ba bước xử lý filter để loại bỏ những pixel trong ảnh không có khả năng trùng với ảnh template truy vấn. Giả sử A là một ảnh xám và T là một ảnh template xám cần matching. Mục tiêu của Círatefi là tìm tất cả các xuất hiện của ảnh template T ở trong A mà ở đó vật thể T trong ảnh A đã bị xoay, phóng to thu nhỏ và dịch chuyển. Dưới đây là miêu tả của phương pháp Círatefi.

## 3. Kỹ thuật correlation coefficient

Phương pháp Círatefi dùng kỹ thuật correlation coefficient hay còn được biết đến là normalized correlation) trong mỗi bước để ước lượng độ tương ứng giữa ảnh template Q tại pixel (x,y) của ảnh gốc A. Gọi  $v$  là vector biểu diễn độ xám trung bình ở các phần xác định của Q và  $w$  là vector biểu diễn độ xám trung bình tương ứng của A(x,y). Phép correlation coefficient được xác định như sau:

$$X(v, w) = \frac{v \cdot w}{|v| \cdot |w|}$$

Ta có  $v \cdot w = |v| \cdot |w| \cdot \cos(v, w)$ . Vậy  $X(v, w) = \cos(v, w)$ . Vậy phép toán trên sẽ cho kết quả là một số thực nằm trong khoảng (-1, 1). Kết quả càng gần 1 thì góc giữa hai vector  $v$  và  $w$  sẽ càng nhỏ nên hai vector sẽ càng gần nhau. Phép toán trên tính toán độ match giữa hai vector.

## 4. Pipeline

### 4.1 Bước một – Circular Sampling Filter (Cífi)

Filter đầu tiên là Cifi (Circular sampling Filter) thực hiện lấy mẫu một tập các đường tròn trên ảnh template cũng như trên ảnh gốc và thực hiện thuật toán correlation coefficient để lấy ra một tập first grade candidate pixels và một tập best matching scales. Cho một ảnh B, định nghĩa circular sampling là phép toán lấy độ xám trung bình của các pixel cách pixel (x, y) một khoảng r:

$$S_B^\Omega(x, y, z) = \frac{1}{2\pi r} \int_0^{2\pi} B(x + r\cos(\theta), y + r\sin(\theta)) d\theta$$

Ở đây ta dùng ký hiệu  $\Omega$  để chỉ phép toán circular sampling. Trong khi cài đặt, vì ta thực hiện trên một ma trận rời rạc nên dùng phép toán lấy xích-ma ( $\Sigma$ ) thay cho phép toán tích phân cho những điểm thỏa mãn khoảng cách từ tâm đến điểm đúng bằng r.

Đối với ảnh template Q và một tập các scale  $s_0, s_1, \dots, s_{n-1}$ , ảnh template Q được resize theo mỗi scale  $s_i$  tạo ra tập các ảnh template được resize  $Q_0, Q_1, \dots, Q_{n-1}$ . Thực hiện phép toán circular sampling trên mỗi ảnh template  $Q_i$  theo một tập  $l$  các bán kính (radii) đã được định nghĩa trước  $r_0, r_1, \dots, r_{l-1}$  tạo ra một mảng 2 chiều với  $n$  hàng (scales) và  $l$  cột (radii)

$$C_Q[i, k] = S_{Q_i}^\Omega(x_0, y_0, r_k), 0 \leq i < n \text{ and } 0 \leq k < l$$

Trong đó  $(x_0, y_0)$  là tâm của Q. Đối với một số scale, một số vòng tròn không nằm vào trong ảnh template được resize. Những vòng tròn này được biểu diễn bởi giá trị -1 và không được dùng trong phép tính toán correlation coefficient.

Cho một ảnh gốc A, ta sẽ xây dựng ma trận ảnh 3 chiều là các ảnh được lấy mẫu với  $l$  bán kính cho mỗi pixel (x, y):

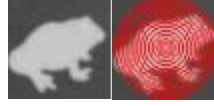
$$C_A[x, y, k] = S_A^\Omega(x, y, r_k), 0 \leq k < l, (x, y) \in \text{domain}(A)$$

Cifi dùng các ma trận  $C_Q$  và  $C_A$  để xác định circular sampling correlation và best matching scale cho mỗi pixel (x, y):

$$X_{A,Q}^\Omega(x, y) = \text{MAX}_{i=0}^{n-1} [X(C_Q[i], C_A[x, y])],$$

Trong đó  $X(C_Q[i], C_A[x, y])$  là phép correlation coefficient giữa hai vector  $C_Q[i]$  và  $C_A[x, y]$ . Pixel (x, y) được phân loại là một “first grade candidate” nếu như  $X_{A,Q}^\Omega(x, y) \geq t_1$  với  $t_1$  là một threshold được định nghĩa trước. Đối với một “first grade candidate”, đại lượng “best matching scale” của nó là đại lượng có giá trị correlation cao nhất:

$$G_{A,Q}^\Omega(x, y) = \text{ARGMAX}_{i=0}^{n-1} [X(C_Q[i], C_A[x, y])]$$



(a) (b)

Figure 1. Ảnh template “frog”. (a) Ảnh template với circular sampling. (b)

#### 4.2 Bước hai – Radial Sampling Filter (Rafi)

Filter thứ hai gọi là Rafi (Radial sampling Filter), tương tự như filter đầu tiên, nó thực hiện lấy mẫu theo một tập các đường thẳng tuyến tính với các góc khác nhau trên ảnh template và ảnh gốc để nâng cấp “second grade candidate” từ các “first grade candidate”. Những pixel không được nâng cấp sẽ bị bỏ đi. Nó gán cho mỗi “second grade candidate” một “best matching rotation angle”. Cho một ảnh B, ta định nghĩa radial sampling là phép toán lấy độ xám trung bình của các pixel nằm trên một đường thẳng tuyến tính gốc tại  $(x, y)$ , độ dài  $\lambda$  và góc nghiêng  $\alpha$ :

$$S_B^\Phi(x, y, \lambda, \alpha) = \frac{1}{\lambda} \int_0^\lambda B(x + t\cos\alpha, y + t\sin\alpha) dt$$

Ở đây ta dùng ký hiệu  $\Phi$  để chỉ phép toán radial sampling. Tương tự như circular sampling. Vì khi cài đặt ta làm việc trên một ma trận rời rạc nên thực hiện phép toán lấy tích-ma ( $\Sigma$ ) thay vì tích phân đối với những điểm thỏa mãn nằm trên đường thẳng tuyến tính gốc tại  $(x, y)$ , độ dài  $\lambda$  và góc nghiêng  $\alpha$ .

Đối với một ảnh template Q và một tập m các góc nghiêng  $\alpha_0, \alpha_1, \dots, \alpha_{m-1}$ , Q thực hiện phép toán radial sampling trên Q với độ dài là  $r_{l-1}$ , là bán kính lớn nhất theo phép lấy mẫu circular sampling, tạo nên một vector với m thuộc tính:

$$R_Q[j] = S_Q^\Phi(x_0, y_0, r_{l-1}, \alpha_j), 0 \leq j < m$$

trong đó  $(x_0, y_0)$  là tâm của ảnh template Q.

Với mỗi “first candidate pixel”  $(x, y)$  thực hiện phép toán radial sampling tại các scale khả dĩ đối với phép toán circular sampling:

$$R_A[x, y, j] = S_A^\Phi(x, y, s_i r_{l-1}, \alpha_j), 0 \leq j < m$$

trong đó  $s_i r_{l-1}$  là bán kính của ảnh template Q đã được scale.

Tại mỗi “first grade candidate” pixel  $(x, y)$ , Rafi dùng vector  $R_A[x, y]$  và  $R_Q$  để xác định radial sampling correlation và best matching angle cho mỗi pixel đó:

$$X_{A,Q}^\Phi(x, y) = \text{MAX}_{j=0}^{m-1} [X(R_A[x, y], \text{cshift}_j(R_Q))]$$

Trong đó  $cshift_j$  là phép toán element-wise rotation  $j$  vị trí của vector tham số. Một “first grade candidate” được nâng cấp thành “second grade candidate” nếu như  $X_{A,Q}^\Phi(x, y) \geq t_2$  với  $t_2$  là một threshold được định nghĩa trước. Đối với một “second grade candidate”, đại lượng “best matching angle” là đại lượng có giá trị correlation lớn nhất:

$$G_{A,Q}^\Phi(x, y) = ARGMAX_{j=0}^{m-1} [X(R_A[x, y], cshift_j(R_Q))]$$

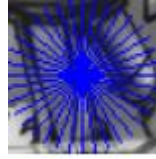


Figure 2. Ảnh template với radial sampling.

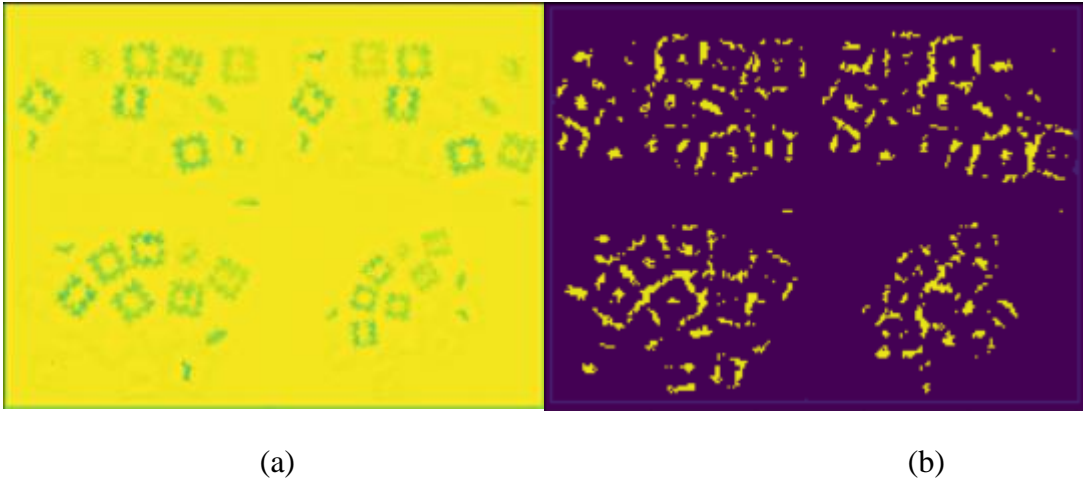
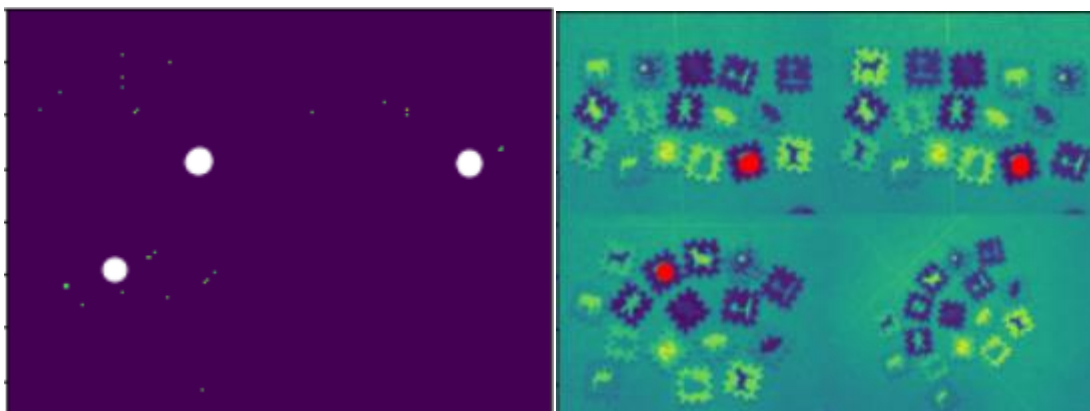


Figure 3. Output của Cifi. (a) Output của Rafi. (b)

### 4.3 Bước ba – Template Matching Filter (Tefi)

Filter thứ ba được gọi là Tefi (Template matching Filter). Với mỗi candidate pixel, tefi xoay và resize ảnh template bởi các scale và góc xoay (là output từ cefi và rafı tương ứng) và thực hiện correlation coefficient ảnh template với candidate pixel tại scale và góc xoay đó. Nếu như kết quả correlation coefficient tại một giá trị lớn hơn giá trị  $t_3$  được định nghĩa trước, có nghĩa là template sẽ tương ứng với vị trí đó ở trên ảnh gốc. Nếu như người dùng biết rằng ảnh template Q chỉ xuất hiện một lần duy nhất ở trên ảnh gốc A, threshold  $t_3$  sẽ không được dùng. Thay vào đó, pixel mà ở đó giá trị correlation coefficient sẽ được chọn.



(a)

(b)

Figure 4. Ouput của Tefi. (a) Kết quả matching tương ứng với template. (b)