

PHƯƠNG PHÁP PHÁT HIỆN ĐỐI TƯỢNG DỰA TRÊN CẤU TRÚC KẾT CẤU NGOẠI HÌNH

Đào Xuân Phúc¹, Nguyễn Thống Nhất¹, Nguyễn Thị Tố Uyên¹
Email: phucdx@hou.edu.vn

Ngày tòa soạn nhận được bài báo: 02/12/2024

Ngày phản biện đánh giá: 11/06/2025

Ngày bài báo được duyệt đăng: 23/06/2025

DOI: 10.59266/houjs.2025.592

Tóm tắt: Bài báo mô tả một phương pháp phát hiện đối tượng dựa trên cấu trúc ngoại hình VTS (Visual Texture Structure). Một số cấu trúc ngoại hình mẫu được lựa chọn để so khớp ảnh, từ đó phát hiện đối tượng trong ảnh thu nhận được. Kết quả nghiên cứu cho thấy phương pháp này có thể phát hiện đối tượng trong ảnh với độ tin cậy cao, có thể phát hiện đối tượng trong điều kiện thông thường, ảnh có nhiễu và trong các điều kiện ánh sáng thay đổi hoặc đối tượng bị che khuất một phần.

Từ khóa: thị giác máy, xử lý ảnh, khớp ảnh, phát hiện đối tượng

I. Đặt vấn đề

Các phương pháp phát hiện đối tượng trong ảnh đã được nghiên cứu từ rất lâu trong lĩnh vực thị giác máy. Phát hiện đối tượng trong ảnh được sử dụng rộng rãi trong nhiều ứng dụng và lĩnh vực khác nhau. Nó đặc biệt hữu dụng trong các bài toán phát hiện, bắt bám thời gian thực. Sự biến đổi vị trí thu được từ quá trình phát hiện mẫu có thể được sử dụng cho nhiều nhiệm vụ khác nhau, ví dụ như trong điều hướng rô bốt trên một dây chuyền sản xuất hoặc là bước đầu tiên trong quá trình bắt bám đối tượng bằng phát hiện đối tượng trước (tracking by detection). Trong hầu hết các trường hợp, quá trình lấy mẫu đã biết trước của đối tượng được thực hiện trên ảnh của đối tượng. Đối tượng có thể được biến đổi

bởi một phép biến đổi nhất định tùy thuộc vào các thiết lập cụ thể, chẳng hạn như phép dịch chuyển, phép biến đổi σ -clit, phép biến đổi tương tự hoặc biến đổi afin. Đã có một số phương pháp được đề xuất để phát hiện đối tượng trong các ảnh dựa trên việc so sánh các giá trị mức xám của điểm ảnh. Các phương pháp này thường không bền vững khi ảnh có nhiễu, ánh sáng thay đổi hay khi đối tượng cần phát hiện bị che khuất. Bài báo này sẽ mô tả việc cải thiện phương pháp phát hiện đối tượng trong ảnh của các phương pháp trên với vật mẫu biết trước dựa trên ảnh mức xám kết hợp cấu trúc kết cấu ngoại hình VTS (Visual Texture Structure) của đối tượng. Kết quả của phương pháp này sẽ được so sánh và kiểm chứng với một số phương pháp phổ biến.

¹ Trường Đại học Mở Hà Nội

II. Phương pháp VTS

Phương pháp đơn giản để phát hiện đối tượng là dựa trên các giá trị mức xám của mẫu và ảnh mục tiêu sử dụng các thuật toán như SAD (Vasanthapriya & cộng sự, 2017). Thuật toán này so sánh trực tiếp sự khác nhau về điểm ảnh của hai ảnh, tổng giá trị tuyệt đối sai khác giữa các điểm ảnh được sử dụng làm tiêu chí đánh giá nếu mẫu có trong ảnh mục tiêu. Theo phương pháp này, ảnh đối tượng O được tìm thấy trong ảnh mục tiêu F nếu giá trị Sad trong (1) là nhỏ nhất.

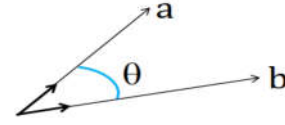
$$Sad(i, j) = \sum_{i=0}^M \sum_{j=0}^N |O(i, j) - F(i + u, j + v)| \quad (1)$$

Trong đó $M \times N$ là kích thước của ảnh F và u, v là các biến chạy dọc theo chiều rộng và chiều cao của ảnh. SAD sử dụng giá trị tuyệt đối nên không nhạy cảm với sự thay đổi hướng của sự khác biệt giữa các điểm ảnh nên nó không chính xác khi ảnh có nhiễu. Tương tự với phương pháp SAD là phương pháp SSD [4], nhưng thay vì tính tổng trị tuyệt đối sự khác nhau giữa các điểm ảnh của O và F thì SSD tính tổng bình phương sự khác nhau giữa các điểm ảnh như trong công thức (2).

$$Ssd(i, j) = \sum_{i=0}^M \sum_{j=0}^N (O(i, j) - F(i + u, j + v))^2 \quad (2)$$

Thuật toán SSD sử dụng bình phương sự khác nhau về giá trị điểm ảnh nên nó nhạy cảm hơn về sự khác biệt giữa các điểm ảnh vì SSD sẽ có sự tăng đột biến khi có sự thay đổi đột ngột trong giá trị của các điểm ảnh. Mặc dù SSD chính xác hơn so với SAD, tuy nhiên cả hai phương pháp này đều không bền vững với ảnh có nhiễu, ánh sáng thay đổi hoặc đối tượng bị che khuất một phần. Fouda và Khan (2015) đã đề xuất phương pháp NCC, phương pháp này cũng dựa trên giá

trị mức xám của điểm ảnh. Thay vì tính toán sự khác nhau về giá trị mức xám giữa các điểm ảnh, NCC được tính theo tính chất tích vô hướng của hai vectơ. Hai vectơ a và b (hình 1) có góc giữa hai vectơ được tính theo công thức (3).



Hình 1. Quan hệ giữa hai vectơ a và b

$$\cos(\theta) = \frac{\langle a, b \rangle}{\|a\| \|b\|} \quad (3)$$

Trong công thức (3), $\cos(\theta)$ đạt giá trị lớn nhất khi hai vectơ a và b cùng hướng và không có sự chênh lệch nào giữa chúng. Xem xét O và F trong (1) là các vectơ thì phép đo tương quan giữa O và F của NCC sẽ dựa vào góc θ như trong công thức (4).

$$Ncc = \sum_i \frac{\langle O_i, F_i \rangle}{\|O\| \|F\|} \quad (4)$$

Ncc có giá trị trong khoảng từ -1 tới +1, Ncc càng lớn thì xác suất mẫu O có mặt trong ảnh mục tiêu F càng lớn. NCC không tính toán dựa trên sự sai khác về giá trị điểm ảnh nên nó khá bền vững với sự thay đổi về ánh sáng so với SAD và SSD. Nghiên cứu của Ahn và Lee (2013) cũng chỉ ra rằng NCC làm việc tốt khi vật mẫu bị quay hoặc che khuất một phần. Tuy nhiên vì NCC vẫn dựa trên giá trị điểm ảnh nên nếu có sự thay đổi sáng lớn hoặc nhiễu thì NCC sẽ cho kết quả kém chính xác. Một số phương pháp phức tạp hơn để phát hiện đối tượng không sử dụng giá trị mức xám của mẫu mà sử dụng cấu trúc cạnh biên của đối tượng để phát hiện (Borgefors, 1988; Rucklidge, 1997). Trong các thuật toán này, cạnh biên của đối tượng cần được phân tách và có thể sử dụng các phép tính

tương quan Hausdorff dựa trên khoảng cách trung bình các cạnh biên của mẫu với các cạnh biên của ảnh (5). Bất lợi của phương pháp này là không bền vững với các trường hợp mẫu bị che khuất.

$$h(O, F) = \|o - f\| \quad (5)$$

Các phương pháp trong Borgefors (1988) và Rucklidge (1997) sử dụng cạnh biên được chuyển sang ảnh nhị phân. Sự biến đổi này làm cho thuật toán phát hiện đối tượng chỉ bền vững khi thay đổi độ sáng trong một phạm vi hẹp. Khi độ tương phản của ảnh giảm xuống tới mức chỉ có một số ít điểm ảnh trên cạnh biên còn có thể phân đoạn được, thì các thuật toán trên trở nên ít hiệu quả. Các phương pháp (1), (2), (3) và (4) đều có điểm chung là không xét đến hướng (gradient) của các cạnh. Nghiên cứu của Olson và Huttenlocher (1997) chỉ ra rằng, việc bỏ qua hướng của các cạnh sẽ dẫn tới việc phát hiện sai mẫu trong ảnh mục tiêu. Việc không xét tới hướng của các cạnh có thể ảnh hưởng đến việc định vị đối tượng, đặc biệt trong trường hợp các cạnh có hướng khác nhau có thể đại diện cho các đặc trưng khác nhau của đối tượng. So sánh với việc có xét tới hướng của các cạnh, việc này có thể cải thiện độ chính xác của quá trình phát hiện, đặc biệt là trong các trường hợp mà hướng của cạnh mang thông tin quan trọng về cấu trúc của đối tượng.

Để tối ưu hóa các vấn đề trên, phương pháp VTS mô tả trong bài báo này kết hợp sử dụng (3) trong khi có tính đến sự thay đổi hướng đặc trưng của ảnh. Bên cạnh đó, VTS sử dụng ảnh mức xám thay vì các ảnh nhị phân trong các tính toán. Cụ thể, ảnh của đối tượng mẫu O và ảnh mục tiêu F cần được chuyển sang ảnh mức xám theo công thức (6) trước khi thực hiện tính hướng của các điểm ảnh.

$$Gray = 0,299R + 0,587G + 0,114B \quad (6)$$

Việc tính hướng để lấy đặc trưng cấu trúc kết cấu ảnh nhờ sử dụng thuật toán Sobel (7) và áp dụng cho cả ảnh đối tượng và ảnh mục tiêu. Gọi G_x và G_y lần lượt là gradient theo trục x và y của điểm ảnh, G_x và G_y được xác định theo công thức (6).

$$\begin{aligned} G_x &= [-1 \ 0 \ 1 \ -2 \ 0 \ 2 \ -1 \ 0 \ 1] * I \\ G_y &= [-1 \ -2 \ -1 \ 0 \ 0 \ 0 \ 1 \ 2 \ 1] * I \end{aligned} \quad (7)$$

Trong đó I là ma trận ảnh gốc (O và F). Khi đó vectơ chỉ phương của điểm ảnh được xác định theo công thức (8).

$$\nabla = \left[\frac{G_x}{\sqrt{G_x^2 + G_y^2}} \quad \frac{G_y}{\sqrt{G_x^2 + G_y^2}} \right] \quad (8)$$

Mẫu của đối tượng cần định vị bao gồm một lớp các điểm ảnh $p_i = (x_i, y_i)^T$ với véc tơ chỉ phương $\nabla_i = (\nabla_{x_i}, \nabla_{y_i})^T$, $i = 1, \dots, N$ (bao gồm các điểm ảnh nền có giá trị bằng 0 và các điểm đặc trưng của vật mẫu có giá trị khác 0). Thông thường mẫu được tạo từ ảnh chứa đối tượng cần định vị, là phần trong ảnh mà đối tượng được định vị. Ảnh chứa mẫu có thể chuyển đổi thành một mô tả trong đó mỗi điểm ảnh có vectơ chỉ phương $D_{x,y} = (u_{x,y}, v_{x,y})^T$. Trong quá trình đối sánh, mẫu đã biến đổi cần được so sánh với ảnh tại vị trí cụ thể. Để phương pháp VTS có thể đối phó với các trường hợp mẫu cần phát hiện bị dịch chuyển hoặc bị xoay so với vị trí lấy mẫu ban đầu, phép biến đổi mẫu được lựa chọn ở đây là phép biến đổi Affin nên mẫu được biến đổi là một lớp các điểm $p'_i = Ap_i$ và vectơ chỉ phương tương ứng là $\nabla'_i = A\nabla_i$, trong đó A (9) là ma trận biến đổi 2D.

$$A = \begin{bmatrix} a & b & t_x & c & d & t_y & 0 & 0 & 1 \end{bmatrix} \quad (9)$$

Áp dụng (3) cho ∇'_i , và D ta được:

$$MS = \sum_{i=1}^N \frac{\langle \nabla'_i, D \rangle}{\|\nabla'_i\| \|D\|} \quad (10)$$

Lấy trung bình MS ở (10) ta được:

$$\underline{MS} = \frac{MS}{N} \quad (11)$$

Việc lấy trung bình là để tính toán một giá trị tổng quát cho độ tương đồng giữa hai hình ảnh, tiện cho việc so sánh. MS ở (11) là phép đo sự tương đồng giữa ảnh F chứa đối tượng cần phát hiện và ảnh đối tượng O của phương pháp VTS, MS càng lớn thì phép định vị càng chính xác, $MS \leq 0$ thì đối tượng cần định vị không có trong ảnh mục tiêu, $MS=1$ thì đối tượng được tìm thấy trong ảnh hết với đối tượng mẫu.

Nhờ việc chuẩn hóa các véc tơ chỉ phương ∇' và I nên MS bất biến với những thay đổi độ sáng vì tất cả các véc tơ được chia tỷ lệ theo độ dài 1. Phép tính (11) đáp ứng tính hiệu quả khi đối tượng bị che khuất một phần, ánh sáng thay đổi hoặc nền có nhiễu. Việc tính tích vô hướng vectơ giữa ảnh mẫu với từng vùng trong ảnh chứa đối tượng mang lại hiệu quả cao bởi các điểm ảnh nền của ảnh mẫu có giá trị bằng 0 sẽ triệt tiêu với các điểm ảnh không thuộc đối tượng cần phát hiện trong ảnh. Nếu đối tượng cần định vị bị che khuất một phần thì cấu trúc kết cấu của đối tượng có mặt ít trong ảnh mục tiêu, khi đó các véc tơ chỉ phương có độ dài nhỏ và như vậy thì MS ở (11) cũng sẽ giảm. Ngược lại nếu đối tượng càng xuất hiện đầy đủ trong ảnh thì MS sẽ lớn.

Giá trị của MS trong (11) có giá trị trong khoảng từ -1 tới 1, giá trị âm thể hiện độ tương phản bị đảo ngược giữa vật được lấy mẫu và vật mẫu trong ảnh mục tiêu. Nếu chỉ lấy giá trị dương để đánh giá kết quả định vị thì trường hợp vật mẫu có trong ảnh mục tiêu bị bỏ qua khi có độ tương phản bị đảo ngược. Vì vậy, để kết quả phát hiện đối tượng không bỏ sót

trường hợp nào, công thức (10) được viết lại như sau:

$$MS = \left| \sum_{i=1}^N \frac{\langle \nabla'_i, D \rangle}{\|\nabla'_i\| \|D\|} \right| \quad (12)$$

Khi đó $\underline{MS}$ trong công thức (11) được viết lại như sau :

$$\underline{MS} = \frac{|MS|}{N} \quad (13)$$

$\underline{MS}$ trong công thức (13) sẽ có giá trị cao nếu các véc tơ chỉ phương của mẫu và ảnh mục tiêu chỉ cùng hướng. $\underline{MS}$ có giá trị bằng 1 nếu đối tượng được tìm thấy hoàn toàn trong ảnh mục tiêu, trong các trường hợp khác $\underline{MS}$ có giá trị nhỏ hơn 1 tương ứng với phần đối tượng được tìm thấy trong ảnh mục tiêu. Căn cứ vào kết quả này, một giá trị ngưỡng có thể được sử dụng để làm tiêu chuẩn xác định đối tượng cần phát hiện có mặt trong ảnh mục tiêu hay không.

III. Đánh giá phương pháp

Phương pháp VTS được tiến hành thực nghiệm với các thiết lập theo tiêu chuẩn công nghiệp, trong đó camera vuông góc với mặt phẳng làm việc và khoảng cách từ camera tới mặt phẳng làm việc là cố định. Nguồn chiếu sáng ổn định và có thể thay đổi được cường độ sáng. Ảnh thu được từ camera được sử dụng có độ phân giải 1280x960.

Vật mẫu và một số tình huống đánh giá phương pháp VTS được đưa ra trong hình 2. Đối tượng được chọn để kiểm tra là vật mẫu như trong hình 2A. Vật mẫu được lấy mẫu nhờ sử dụng thuật toán tiền xử lý (Sobel & Feldman, 1968) áp dụng cho vùng hình chữ nhật bao quanh vật mẫu. Vật mẫu có cấu trúc kết cấu ngoại hình như trong hình 2B. Như đã mô tả ở trên, trong quá trình đối sánh mẫu và ảnh mục tiêu, ảnh mục tiêu cũng được áp dụng các thuật toán tiền xử lý (Sobel & Feldman, 1968).

Các tiêu chí đánh giá phương pháp VTS bao gồm: độ tin cậy khi vật mẫu bị che khuất, vật mẫu bị quay theo hướng bất kỳ, ảnh bị nhiễu, ánh sáng thay đổi. Kết quả của phương pháp VTS được so sánh với các phương pháp trong nghiên cứu của Vasanthapriya và cộng sự (2017), Gonzalez và Woods, (2018), Fouda và Khan (2015). Các kết quả đánh giá phương pháp VTS được đưa ra như trong các bảng dưới đây trong đó: R_s là kết quả định vị đối tượng ($R_s = \blacksquare$ là định vị được đối tượng, $R_s = \square$ là không định vị được đối tượng); ms là độ tin cậy của kết quả định vị đối tượng có giá trị trong khoảng từ 0 tới 1; t là thời gian mà phương pháp phát hiện ra đối tượng trong ảnh có đơn vị tính bằng mili giây.

Đánh giá VTS khi vật mẫu bị che khuất: Vật mẫu được che khuất bởi một vật bằng kim loại (hình 2C, vật bị che $\frac{1}{2}$ thân).

Bảng 1. Kết quả định vị đối tượng của các phương pháp khi vật mẫu bị che khuất

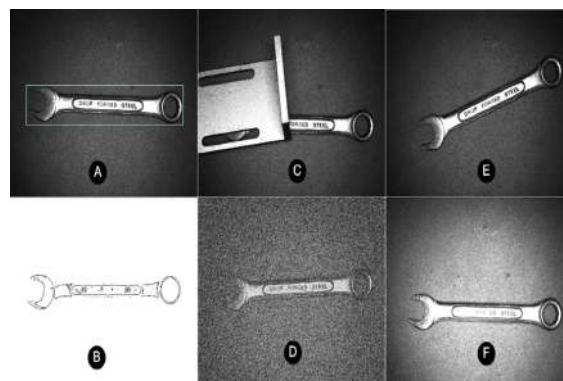
Tỷ lệ bị khuất	SSD			NCC			VTS		
	R_s	ms	t	R_s	ms	t	R_s	ms	t
1*	$\blacksquare$	0,98	5,3	$\blacksquare$	0,99	6,5	$\blacksquare$	0,99	5,4
1/8	$\square$			$\blacksquare$	0,73	6,6	$\blacksquare$	0,91	6,3
1/4	$\square$			$\square$			$\blacksquare$	0,84	6,6
1/2	$\square$			$\square$			$\blacksquare$	0,58	6,9
3/4	$\square$			$\square$			$\square$		

Đánh giá VTS khi vật mẫu bị quay: Vật mẫu được kiểm tra khi quay vật mẫu trong khoảng từ 0 tới 360 độ (hình 2E, vật bị quay 30 độ). Việc thực hiện biến đổi afin trong VTS và các thuật toán so sánh

Bảng 2. Kết quả định vị đối tượng của các phương pháp khi vật mẫu bị quay

Góc quay	SSD			NCC			VTS		
	R_s	ms	t	R_s	ms	t	R_s	ms	t
0*	$\blacksquare$	0,99	5,5	$\blacksquare$	0,98	6,6	$\blacksquare$	0,99	5,4
10	$\blacksquare$	0,93	5,6	$\blacksquare$	0,96	6,4	$\blacksquare$	0,97	5,7
20	$\square$	0,87	5,7	$\blacksquare$	0,93	6,4	$\blacksquare$	0,97	6,0
30	$\square$	0,85	5,7	$\blacksquare$	0,90	6,7	$\blacksquare$	0,96	6,1
50	$\square$	0,81	5,8	$\blacksquare$	0,88	6,9	$\blacksquare$	0,96	6,3
150	$\square$	0,76	6,8	$\blacksquare$	0,83	7,1	$\blacksquare$	0,95	7,2
270	$\square$	0,73	7,0	$\blacksquare$	0,80	7,5	$\blacksquare$	0,96	6,4

Phương pháp VTS cho thấy sự bền vững, hệ số ms thể hiện rõ sự phù hợp giữa độ chính xác và tỷ lệ bị che khuất của vật mẫu như đã trình bày trong nội dung nghiên cứu (hệ số che khuất lớn thì độ chính xác giảm – bảng 1). Phương pháp SSD phụ thuộc chính vào sự sai khác trực tiếp giữa hai điểm ảnh nên cho kết quả kém khi vật bị che khuất.



Hình 2. Vật mẫu và một số tình huống đánh giá phương pháp VTS

bên cạnh việc ảnh không chịu tác động của nhiễu và các thay đổi khác nên nhìn chung các thuật toán đều cho kết quả tốt khi có thể định vị vật mẫu quay từ 0 tới 360 độ với độ chính xác cao (bảng 2).

Đánh giá VTS khi ảnh bị nhiễu: Nhiễu ở đây là nhiễu hạt được tạo theo phân bố Gaussian (Shin, Kim, & Kim, 2010) (hình 2D, ảnh với hệ số nhiễu bằng 120). Cả SSD và NCC đều phụ thuộc vào cường độ sáng của điểm ảnh mà

không xét tới hướng của cấu trúc kết cấu hình dạng của đối tượng nên bị ảnh hưởng lớn khi có nhiễu. Kết quả cho thấy phương pháp VTS có thể tìm thấy vật mẫu ngay cả khi ảnh có nền nhiễu cao (bảng 3).

Bảng 3. Kết quả định vị đối tượng của các phương pháp khi vật mẫu bị nhiễu

Hệ số nhiễu	SSD			NCC			VTS		
	Rs	ms	t	Rs	ms	t	Rs	ms	t
0*	■	0,99	5,2	■	0,99	6,4	■	0,98	5,3
30	■	0,76	5,5	■	0,87	6,8	■	0,88	5,5
50	■	0,52	5,6	■	0,73	6,8	■	0,81	5,5
90	□			■	0,61	7,0	■	0,64	5,8
120	□			□			■	0,55	5,8
150	□			□			■	0,53	5,6

Đánh giá VTS khi ánh sáng thay đổi: Khi ảnh chứa vật mẫu có sự thay đổi ánh sáng, vật mẫu làm bằng kim loại có độ tương phản với ánh sáng lớn khi cường độ ánh sáng lớn, kết quả cho thấy phương

pháp VTS cho kết quả tốt so với các phương pháp được so sánh (bảng 4). Khi tính chất của điểm ảnh bị thay đổi, phương pháp SSD sẽ suy giảm do phụ thuộc vào giá trị của điểm ảnh.

Bảng 4. Kết quả định vị đối tượng của các phương pháp khi có thay đổi của ánh sáng

Hệ số ánh sáng	SSD			NCC			VTS		
	Rs	ms	t	Rs	ms	t	Rs	ms	t
7	□			□			□		
8	□			■	0,61	6,4	■	0,67	5,9
15	■	0,73	5,8	■	0,85	6,5	■	0,87	5,9
25	■	0,81	5,9	■	0,90	6,3	■	0,93	5,8
36*	■	0,98	5,8	■	0,99	6,3	■	0,98	5,5
45	□			■	0,81	6,3	■	0,83	5,9
75	□			□			■	0,53	5,8
76	□			□			□		

Trong thực nghiệm này, vật mẫu bằng kim loại có bề mặt trơn và độ phản xạ ánh sáng cao nên độ khó trong trích xuất đặc trưng cao. Vì thế, phương pháp SSD đơn giản cho kết quả kém trong cả bốn bài đánh giá. Phương pháp NCC cho kết quả tốt hơn khi ánh sáng thay đổi và vật mẫu bị quay ở một phạm vi nhỏ, không cho kết quả tốt trong các tiêu chí đánh giá còn lại.

IV. Kết luận

Bài báo đã mô tả việc cải thiện phương pháp phát hiện đối tượng trong ảnh của các phương pháp dựa trên so sánh giá

trị mức xám của điểm ảnh với vật mẫu biết trước dựa trên ảnh mức xám kết hợp cấu trúc kết cấu ngoại hình của đối tượng. Kết quả cho thấy phương pháp này cho kết quả tốt khi ảnh mục tiêu có nền nhiễu cao, ánh sáng thay đổi, vật mẫu bị thay đổi hướng và vị trí hay vật mẫu bị che khuất một phần. Kết quả của phương pháp đã được ứng dụng trong các ứng dụng điều hướng cánh tay rô bốt trên các dây chuyền sản xuất công nghiệp. Quá trình đánh giá phương pháp VTS được thực hiện trong các thiết lập tiêu chuẩn công nghiệp, sự thay đổi các điều kiện kiểm tra đều được thực hiện

trong một mặt phẳng làm việc vuông góc với camera mà không có sự biến dạng 3D của vật mẫu. Các nghiên cứu tiếp theo sẽ tập trung mở rộng phương pháp này khi có xét tới các biến đổi phối cảnh của đối tượng trong một trường hợp tổng quát hơn.

Tài liệu tham khảo

- [1]. Ahn, J. H., Lee, J. Y. (2013). *Robust Object Detection under Occlusion and Partial Occlusion*, Robust Object Detection under Occlusion and Partial Occlusion, Volume: 22, Issue: 1.
- [2]. Archana, J., & Aishwarya, P. (2016). A Review on the Image Sharpening Algorithms Using Unsharp Masking, *Int. J. Eng. Sci. Comput.*, 6, 8729–8733.
- [3]. Ballard, D. H. (1981). Generalizing the Hough transform to detect arbitrary shapes, *Pattern Recognition*.
- [4]. Borgefors, G. (1988). Hierarchical chamfer matching: A parametric edge matching algorithm, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(6), 849–865.
- [5]. Brown, L. G. (1992). A Survey of Image Registration Techniques, *ACM Computing Surveys (CSUR)*, Vol. 24, No. 4, 1992.
- [6]. Fouda, Y. M., & Khan, A. R. (2015). Normalize Cross Correlation Algorithm in Pattern Matching Based on 1-D Information Vector, *Trends in Applied Sciences Research*, 10, 195-206.
- [7]. Gonzalez, R. C., & Woods, R. E. (2018). *Digital Image Processing*, 4th ed., Pearson.
- [8]. Olson, C. F., Huttenlocher, D. P. (1997). Automatic Target Recognition by Matching Oriented Edge Pixels, *IEEE TRANS. ON IMAGE PROCESSING*, vol. 6, no. 1, 1997.
- [9]. Rucklidge, W. J. (1997). Efficiently locating objects using the Hausdorff distance, *International Journal of Computer Vision*, 24(3), 251–270.
- [10]. Shin, J., Kim, J., & Kim, K. (2010). A Noise Generation Model for Gaussian Noise in Digital Images, *International Journal of Imaging Systems and Technology*.
- [11]. Sobel, I., & Feldman, G. (1968). *A 3x3 Isotropic Gradient Operator for Image Processing*, Stanford Artificial Intelligence Project.
- [12]. Tanimoto, S. L. (1981). Template matching in pyramids, *Computer Graphics and Image Processing*, 16, 356–369, 1981.
- [13]. Vasanthapriya, K, Vaishnavi, P., Nithya Lakshmi.A, “Implementation of SAD algorithm for motion estimation”, *IJARIE-ISSN(O)-2395-4396*, Vol-3 Issue-2 2017.

A METHOD FOR OBJECT DETECTION IN IMAGES BASED ON VISUAL TEXTURE STRUCTURE

Dao Xuan Phuc², Nguyen Thong Nhat², Nguyen Thi To Uyen²

Abstract: This article presents a method for object detection based on visual texture structure (VTS). A number of sample objects are selected to match the images, thereby detecting objects in the received images. Research results show that this method detects objects in images with high reliability, can detect in normal conditions, images with noise and in changing lighting conditions or partial occlusion.

Keywords: machine vision, image processing, object detection, visual texture structure

² Hanoi Open University