3. SURF

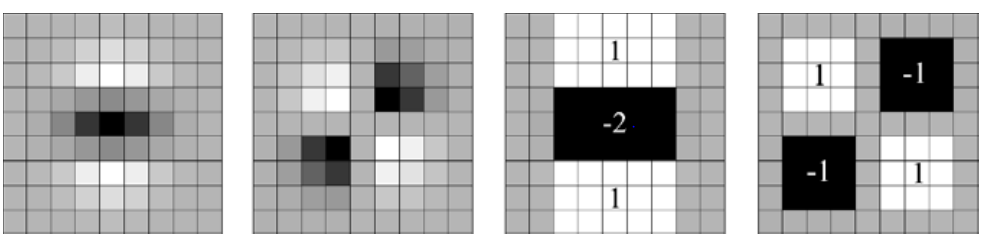
SURF, do Herbert Bay đề xuất, là một thuật toán lấy ý tưởng cơ bản từ SIFT, nhưng có những khác biệt nhất định. Giống như SIFT, SURF cũng có 2 tính năng là phát hiện điểm chú ý và trích xuất mô tả từ các điểm đó. Các đặc tả của SURF cũng có thể được dùng trong các mục đích tương tự SIFT như so khớp ảnh…

3.1 Phát hiện điểm chú ý

SURF sử dụng phương pháp tính xấp xỉ đạo hàm bậc 2 của Gaussian bằng bộ lọc 9×9 để tính nhanh định thức của ma trận Hessian của hàm Gaussian đối với từng điểm ảnh. Ma trận Hessian của ảnh I tại điểm ảnh x và scale δ là

, Lxx là đạo hàm bậc 2 của hàm Gaussian, tương tự với Lxy…

Việc tính toán định thức của ma trận trên tiêu tốn khá nhiều tài nguyên và thời gian, do đó Bay và cộng sự đã đề xuất một cách tính xấp xỉ bằng cách đưa các đạo hàm về các bộ lọc số nguyên 9×9.



Hình : 2 hình bên trái lần lượt là biểu diễn các đạo hàm Gaussian(Lyy và Lxy), 2 hình bên phải là các xấp xỉ của chúng.

Sau khi tính toán xong, các định thức sẽ được tạo thành một bảng định thức. Sau đó kích thước của bộ lọc sẽ đươc tăng lên và thuật toán lặp lại vài lần để tạo thành 1 octave. Khác với sub-sampling của SIFT, SURF tăng kích thước các bộ lọc thay vì giảm kích cỡ ảnh để xử lý, do đó ảnh vẫn được giữ nguyên kích thước.

Với mỗi một scale của một octave, một khu vực lân cận 3×3×3 được xét với điểm trung tâm của khối sẽ là điểm được chọn nếu giá trị định thức tại ô đó là giá trị cực đai so với các điểm còn lại trong khu vực lân cận. Điểm này sau đó sẽ được nội suy để tìm ra vị trí điểm chú ý chính xác.

3.2 Trích xuất mô tả

3.2.1 Định hướng điểm

Để có được tính bất biến về góc xoay, định hướng của điểm chú ý cần được tim. Các phản hồi sóng Haar trên cả 2 phương ngang và dọc trong khu vực tròn xung quanh điểm chú ý với bán kính 6s (s là scale tại điểm chú ý đang xét) được tính toán và được đặt trọng số bởi một hàm Gaussian tại điểm chú ý đang xét, sau đó được chuyển thành các điểm trong hệ tọa độ 2 chiều, với phản hồi ngang là hoành độ và phản hồi dọc là tung độ.

Hướng chủ đạo sẽ được ước lượng bằng cách tính tổng của các phản hồi trong cửa sổ quay với độ lớn π/3, phản hổi dọc và ngang trong cửa sổ sẽ được cộng lại tương ứng, sau đó một vector hướng sẽ được dựng lên từ 2 tổng thu được. Vector lớn nhất sẽ là vector hướng chủ đạo. Kích thước của cửa sổ trượt là một tham số phải được lựa chọn cẩn thận để đạt được sự cân bằng mong muốn giữa độ mạnh và độ phân giải góc.

3.2.2 Mô tả dựa trên tổng số phản hồi sóng Haar

Để mô tả khu vực xung quanh điểm, một khu vực hình vuông được trích xuất, tập trung vào điểm chú ý và được định hướng dọc theo hướng như đã chọn ở trên. Kích thước của cửa sổ này là 20s.

Vùng chú ý được chia thành các tiểu vùng vuông nhỏ hơn 4×4 và đối với mỗi vùng, các phản ứng sóng con Haar được trích xuất tại vùng 5×5 các điểm mẫu. Các phản hồi được cân bằng với một hàm Gaussian (để mang lại sự mạnh mẽ hơn cho các biến dạng, tiếng ồn và biến đổi).