

PERBANDINGAN ALGORITMA *SCALE INVARIANT FEATURE TRANSFORM (SIFT)* DAN *SPEEDED UP ROBUST FEATURE (SURF)* PADA PENGENALAN OBJEK BERBASIS MATLAB

Taufik Hidayat¹, Yassir², Syamsul³

^{1,2,3}Prodi Teknologi Rekayasa Jaringan Telekomunikasi
Jurusan Teknik Elektro Politeknik Negeri Lhokseumawe
Email: thidayat170@gmail.com, yassirajalil@gmail.com, syamsul0466@gmail.com

ABSTRAK

Pengenalan objek banyak digunakan dalam industri untuk keperluan inspeksi, registrasi, dan manipulasi. Kesulitan masalah pengenalan objek sebagian besar disebabkan oleh kurangnya keberhasilan dalam menemukan fitur gambar tersebut. Penelitian ini menggunakan 2 metode untuk deteksi dan mendeskripsi fitur, diantaranya *Scale Invariant Feature Transform (SIFT)* dan *Speeded Up Robust Feature (SURF)*. Kedua metode deskriptor fitur ini tangguh pada perubahan skala, keburaman, rotasi, perubahan iluminasi, dan transformasi affine pada objek yang terdapat di sebuah citra. SIFT adalah algoritma yang digunakan untuk mengekstrak fitur dari gambar. SURF adalah algoritma yang efisien, sama seperti SIFT tetapi dengan kompleksitas komputasi lebih rendah. Dalam percobaan ini, jumlah fitur yang terdeteksi oleh SIFT relatif lebih banyak dengan rata-rata presentase data benar sebesar 100% pada *dataset* pertama dan 97.83% pada *dataset* kedua, sedangkan jumlah fitur yang dideteksi oleh SURF sebesar 81.8% untuk *dataset* pertama dan 88.53% untuk *dataset* kedua. Juga kecepatan komputasi SURF relatif lebih cepat daripada SIFT. Diketahui kecepatan rata-rata algoritma SIFT dalam mencari *matched feature* sebesar 72.545 detik pada *range* resolusi dari 756x1008 (*dataset* pertama) sampai 960x540 (*dataset* kedua), dan SURF sebesar 6.58 detik. Sehingga dapat disimpulkan bahwa dalam hal akurasi SIFT lebih baik daripada SURF, namun dalam hal kecepatan komputasi SURF lebih baik daripada SIFT. Oleh sebab itu SURF lebih cocok digunakan untuk keperluan *object tracking* pada citra bergerak (video).

Kata-kata kunci: *SIFT, SURF, Ekstraksi Fitur, Pendeteksian dan deskripsi fitur, Pengenalan objek*

I. PENDAHULUAN

Ilmu dan teknologi pengolahan citra digital merupakan ilmu yang bermanfaat dalam kehidupan. Pengolahan citra digital merupakan pemrosesan gambar melalui komputer untuk mendapatkan informasi tertentu. Beberapa manfaat pengolahan citra digital ialah untuk mendapatkan informasi yang berupa objek apa saja yang terdapat dalam sebuah gambar atau pengenalan objek yang apabila dengan mata sulit dikenali.

Pengenalan objek banyak digunakan dalam industri untuk keperluan inspeksi, registrasi, dan manipulasi. Kesulitan masalah pengenalan objek sebagian besar disebabkan oleh kurangnya keberhasilan dalam menemukan fitur gambar tersebut. Saat ini, fitur citra digunakan untuk pengenalan dan pencocokan dari dua citra. Pada penelitian ini, penulis mengenalkan 2 algoritma *Feature Extraction and Matching* yaitu SIFT dan SURF.

Scale Invariant Feature Transform (SIFT) adalah algoritma yang paling banyak digunakan di antara deskriptor fitur lokal. Algoritma ini pertama kali diusulkan pada tahun 1999 dan disempurnakan pada tahun 2004. SIFT tidak hanya memiliki invariansi perubahan seperti skala, rotasi, dan sudut pandang tertentu dan perubahan iluminasi. sehingga telah

diterapkan secara luas dalam pengenalan objek, dan rekonstruksi 3D. *Speeded Up Robust Features (SURF)* mirip dengan SIFT. SURF memiliki kemiripan dengan SIFT dalam kinerja dan mengurangi kompleksitas komputasi dalam pencocokan gambar. SURF menggunakan respon Haar wavelet untuk mendekati operasi gradien seperti SIFT dan juga teknik grafik integral.

Dalam hal ini penulis melakukan perbandingan pengenalan objek dari 2 algoritma yang berbeda yakni, SIFT dan SURF. Perbandingan yang dilakukan akan memberikan parameter *output* berupa *Detected Feature Points* antara train image dan test image, *Matching Feature Point*, dan *Time taken for Feature Matching*. Penelitian ini akan membuktikan perbedaan algoritma dalam pencocokan yang lebih akurat dan waktu komputasi.

II. TINJAUAN PUSTAKA

A. Citra Digital

Citra digital adalah sebuah fungsi 2D, $f(x,y)$, yang merupakan fungsi intensitas cahaya, dimana nilai x dan y merupakan koordinat spasial dan nilai fungsi di setiap titik (x,y) merupakan tingkat keabuan citra pada model *grayscale* atau RGB dalam model RGB.

Citra digital dinyatakan dengan sebuah matriks dimana baris dan kolomnya menyatakan suatu titik pada citra tersebut dan elemen matriksnya (yang disebut sebagai elemen gambar atau piksel) menyatakan tingkat keabuan pada titik tersebut. Matriks dari citra digital berukuran NxM (tinggi x lebar), dimana:

$$\begin{aligned}
 N &= \text{jumlah baris } 0 \leq y \leq N - 1 \\
 M &= \text{jumlah kolom } 0 \leq x \leq M - 1 \\
 L &= \text{derajat keabuan } 0 = f(x,y) = L - 1
 \end{aligned}$$

Berikut ini adalah gambaran matriks dari citra digital:

$$f(x,y) \approx \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,M-1) \\ f(1,0) & f(1,1) & \dots & f(1,M-1) \\ \vdots & \vdots & \vdots & \vdots \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1,M-1) \end{bmatrix} \quad (1)$$

Dimana indeks baris (x) dan indeks kolom (y) menyatakan suatu koordinat titik pada citra, sedangkan f(x,y) merupakan intensitas (derajat keabuan) pada titik (x,y).

B. Scale Invariant Feature Transform (SIFT)

SIFT adalah algoritma yang diusulkan oleh David G Lowe pada tahun 2001. SIFT memiliki invarian skala, rotasi, dan translasi. SIFT tahan terhadap perubahan iluminasi. Algoritma ini digunakan untuk pengenalan objek berdasarkan pencocokan fitur. Ide utama dari algoritma SIFT adalah untuk menemukan lokal ekstrema dalam skala-ruang gaussian, dan kemudian mengambil titik lokal ekstrema tersebut untuk dijadikan titik fitur yang stabil. Kemudian, karakteristik lokal dari citra diekstraksi di sekitar setiap titik fitur yang stabil dan membentuk deskriptor lokal yang digunakan untuk pencocokan. Proses implementasi SIFT terdiri dari empat tahap [2], yaitu :

1. Scale-space Extrema Detection

Menurut Koenderink (1984) dan Lindeberg (1994) satu-satunya kernel skala-ruang yang adalah fungsi Gaussian [1]. Oleh karena itu, skala-ruang suatu citra didefinisikan sebagai fungsi $L(x, y, \sigma)$ yang dihasilkan dari konvolusi Gaussian skala variabel $G(x, y, \sigma)$, dengan input citra $I(x, y)$.

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y) \quad (2)$$

$$G(x,y,\sigma) = \left(\frac{1}{2\pi\sigma^2}\right) e^{-(x^2+y^2)/2\sigma^2} \quad (3)$$

Pendeteksian lokasi *keypoint* yang stabil dalam skala-ruang dilakukan menggunakan *scale-space extrema* dalam fungsi *difference-of-Gaussian* yang dikonvolusikan dengan citra, $D(x, y, \sigma)$ yang dapat dihitung dari selisih 2 skala terdekat melalui sebuah konstanta faktor k [2].

$$\begin{aligned}
 D(x,y,\sigma) &= (G(x,y,k\sigma) - G(x,y,\sigma)) * I(x,y) \\
 &= L(x,y,k\sigma) - L(x,y,\sigma)
 \end{aligned} \quad (4)$$

2. Keypoint Localization

Setelah calon *keypoint* telah ditemukan dengan membandingkan sebuah piksel ke piksel terdekat, langkah berikutnya adalah untuk melakukan rincian yang sesuai pada data terdekat untuk lokasi, skala dan rasio kurva. Informasi ini memungkinkan titik ditolak yang mempunyai rendah kontras (dan karenanya sensitif terhadap *noise*) atau lokasi yang buruk sepanjang tepi. Implementasi awal pendekatan *keypoint* terletak di lokasi dan skala dari titik tengah sampel. Pendekatannya menggunakan perluasan Taylor (hingga kuadrat) pada fungsi skala-ruang $D(x, y, \sigma)$ digeser sehingga yang asli merupakan titik sampel. besarnya gradien, $m(x, y)$, dan orientasi, (x, y) , dikomputasi menggunakan perbedaan piksel:

3. Orientation Assignment

Dengan menetapkan orientasi yang konsisten untuk setiap *keypoint* berdasarkan sifat lokal citra, deskripsi *keypoint* dapat diwakili terhadap orientasi ini karena itu mencapai variasi untuk rotasi citra. Mengikuti penelitian dengan sejumlah pendekatan untuk menetapkan orientasi lokal, telah ditemukan untuk memberikan hasil yang paling stabil. Skala *keypoint* digunakan untuk memilih citra Gaussian yang baik, L, dengan skala terdekat, sehingga semua komputasi dilakukan dengan cara skala invarian. Untuk setiap sampel citra, $L(x, y)$, pada skala ini, besarnya gradien, $m(x, y)$, dan orientasi, (x, y) , dikomputasi menggunakan perbedaan piksel:

$$\begin{aligned}
 m(x,y) &= \sqrt{[L(x+1,y) - L(x-1,y)]^2 + [L(x,y+1) - L(x,y-1)]^2} \\
 \theta(x,y) &= \tan^{-1} \frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}
 \end{aligned} \quad (5)$$

4. Keypoint Descriptor

Besarnya gradien dan orientasi sampel di sekitar lokasi *keypoint*, menggunakan skala *keypoint* untuk memilih tingkat Gaussian untuk citra yang tidak jelas. Untuk mencapai orientasi invarian, koordinat dan deskripsi orientasi gradien diputar terhadap orientasi *keypoint*. Sebuah deskriptor *keypoint* dibuat dengan terlebih dahulu menghitung besaran gradien dan orientasi pada setiap titik sampel citra di wilayah sekitar lokasi *keypoint*, sampel-sampel ini kemudian diakumulasikan ke dalam histogram orientasi yang meringkas konten di atas subwilayah 4x4, ini menunjukkan *array* deskriptor 2x2 yang dihitung dari kumpulan sampel 8x8.

C. *Speeded Up Robust Features (SURF)*

SURF adalah algoritma yang diusulkan oleh Bay dan rekan-rekannya pada tahun 2007. SURF mendeteksi *keypoint* citra dan menghasilkan deskriptor. SURF memperoleh *keypoints* berdasarkan matriks Hessian [3]. Berdasarkan gambar integral dan matriks Hessian, deskriptor *keypoints* dideteksi menggunakan algoritma SURF. *Keypoints* dari dua gambar dicocokkan untuk pengenalan. Proses implementasi SURF terdiri dari tiga tahap, yaitu :

1. *Keypoint Detection*

Algoritma SURF menggunakan matriks Hessian untuk mengekstrak titik fitur, sehingga matriks Hessian merupakan inti dari algoritma SURF. Dengan asumsi bahwa gambar adalah I, maka perlu dilakukan *filtering* Gauss sebelum membangun matriks Hessian

$$H(X, \sigma) = \begin{bmatrix} Lxx(x, \sigma) & Lxy(x, \sigma) \\ Lxy(x, \sigma) & Lyy(x, \sigma) \end{bmatrix} \quad (6)$$

Di sini, $Lxx(x, \sigma)$ mengacu pada konvolusi dari Gaussian $\partial^2 / \partial x^2 g(\sigma)$ Turunan orde kedua dengan bayangan I di titik X, dimana $g(\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}$. Demikian pula definisi untuk $Lxy(x, \sigma)$ dan $Lyy(x, \sigma)$.

2. *Orientation Assignment*

Orientasi *keypoint* yang terdeteksi ditentukan dengan menggunakan Haar wavelet. Respons dari Haar wavelet dihitung dalam arah x dan y dalam lingkungan melingkar dengan radius 6σ di sekitar titik yang terdeteksi untuk sekumpulan piksel. Respons dari wavelet Haar dijumlahkan dan dihitung dalam jendela orientasi geser berukuran $\pi/3$ untuk menentukan orientasi dominan. Orientasi lokal dapat diperoleh dengan menjumlahkan semua respons x dan y dalam jendela orientasi di setiap lokasi. Dengan mempertimbangkan vektor terpanjang di antara semua jendela, orientasi interest point dapat ditentukan..

3. *SURF Descriptors*

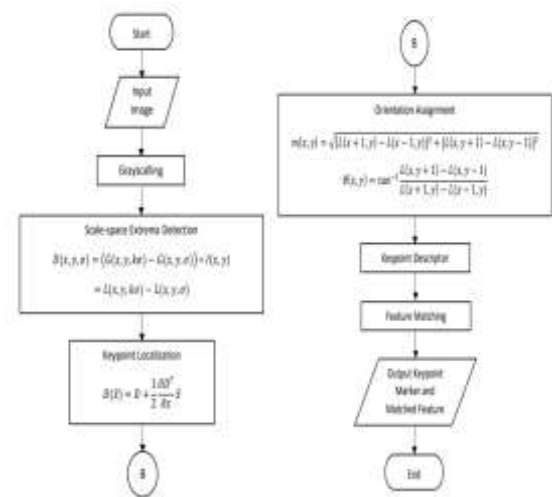
Memberikan deskripsi fitur yang unik dan kuat adalah tujuan utama deskriptor SURF. Berdasarkan area di sekitar *interest point*, deskriptor dapat diperoleh. SURF deskriptor dapat dihitung berdasarkan respon dari Haar wavelet dan dengan citra integral. Mula-mula, di sekitar *keypoint* dibuatlah sebuah daerah persegi. Daerah jendela ini dibagi menjadi sub-daerah yang lebih kecil yaitu $4*4$ persegi, yang dijelaskan oleh nilai-nilai respons wavelet dalam arah x dan y. Ini menjaga informasi spasial yang penting. *Interest areas* dibobot dengan Gaussian yang berpusat pada *interest point* untuk memberikan ketahanan terhadap deformasi dan translasi. Pada $5S*5S$,

dihitung respon Haar wavelet untuk setiap *keypoint*. Oleh karena itu $64D$ ($4*4*4 = 64$) adalah vektor deskriptor dari algoritma SURF. Deskriptor SURF adalah invarian pencahayaan karena respon Haar adalah invarian pencahayaan

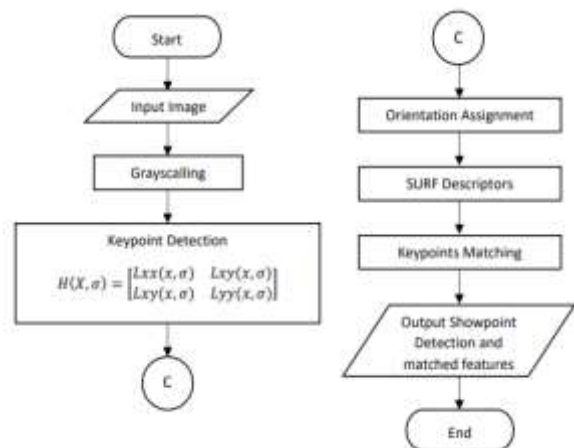
III. METODOLOGI

A. Diagram Alir

Berikut teknik pengolahan data yang dilakukan menggunakan algoritma *Scale Invariant Feature Transform (SIFT)* dan algoritma *Speeded Up Robust Feature (SURF)* ditunjukkan pada Gambar 1 dan Gambar 2.



Gbr 1 Diagram Alir Algoritma SIFT



Gbr 2 Diagram Alir Algoritma SURF

B. Metode Simulasi

Pada penelitian ini menggunakan Matlab sebagai visualisasi gambar yang diproses menggunakan algoritma SIFT dan SURF. *Output* dari masing-masing

algoritma tersebut kemudian dianalisis sehingga diketahui perbedaan dari kedua algoritma.

C. Metode Analisis

Metode analisis ditinjau dari *output* data *Detected Feature Points*, *Matching Feature Point*, dan *Time taken for Feature Matching*. *Detected Feature Points* adalah jumlah *keypoints* yang dideteksi pada citra yang berfungsi untuk mendeskripsikan citra. *Matching Feature Point* adalah jumlah *keypoint* yang cocok antara *test image* dan *training image*. Dan *time taken for feature matching* adalah waktu yang dibutuhkan algoritma untuk memproses pencocokan citra pada pengenalan objek.

Output Data tersebut dicatat dalam tabel untuk memperoleh data kuantitatif sehingga dapat dianalisis kelebihan dan kekurangan antara algoritma SIFT dan SURF. Kinerja algoritma dapat diketahui berdasarkan waktu yang dibutuhkan algoritma dalam memproses pencocokan citra. Dan keakuratan dapat diketahui berdasarkan banyaknya *keypoint* yang dideteksi pada citra

IV. HASIL DAN PEMBAHASAN

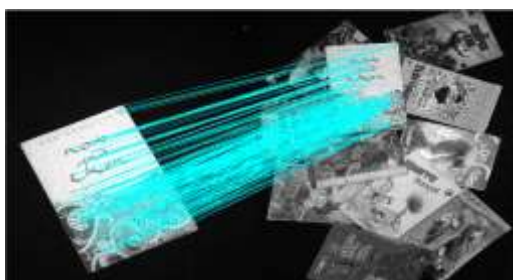
Dalam penelitian ini dilakukan analisis sifat dari dua algoritma yang dapat mendeteksi dan mendeskripsi titik fitur citra 2D untuk keperluan *object recognition*. Pengujian ini dilakukan dengan format gambar JPEG. Citra yang digunakan sebagai *training image* dan *test image* dibagi menjadi dua *dataset*.

A. Pengujian *Dataset* Pertama

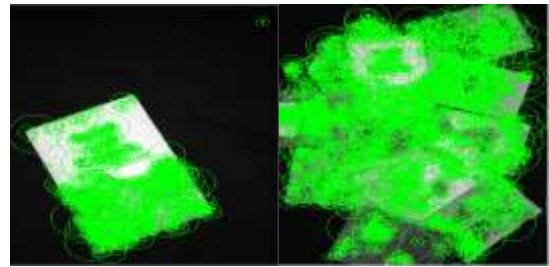
Dataset Pertama yang diuji berupa foto dengan resolusi 756 x 1008. Kedua algoritma digunakan untuk mendeteksi titik fitur yang kuat yang terdapat pada citra. Hasil pengujian dapat dilihat pada gambar 3 sampai gambar 6, dan tabel 1.



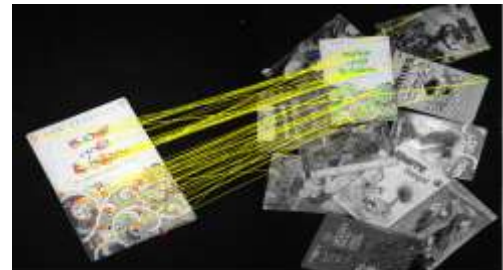
Gbr 3 Fitur yang terdeteksi oleh SIFT



Gbr 4 Fitur yang cocok oleh SIFT



Gbr 5 Fitur yang terdeteksi oleh SURF



Gbr 6 6 Fitur yang cocok oleh SURF

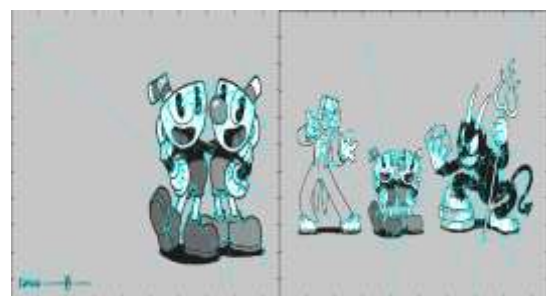
Gambar 3 menunjukkan hasil fitur yang terdeteksi dari *train image* dan *test image* menggunakan algoritma SIFT. Sedangkan Gambar 5 menggunakan algoritma SURF.

TABEL I
Hasil Pengujian *Dataset* Pertama

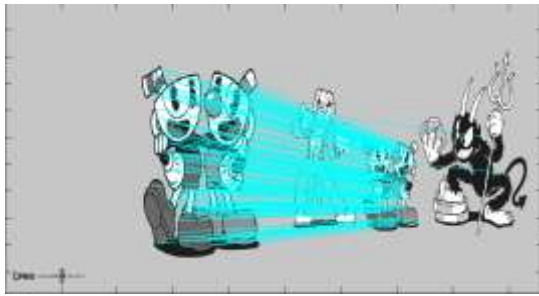
Deskriptor Fitur	Titik Fitur Terdeteksi		Matching Titik Fitur	Waktu yang dibutuhkan untuk Matching Fitur
	Training image	Test image		
SIFT	1705	5696	300	130.19 s
SURF	892	2540	63	9.55

B. Pengujian *Dataset* Kedua

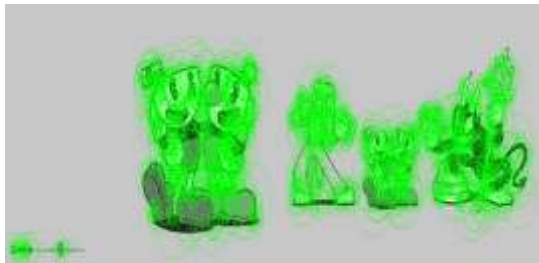
Dataset Kedua berupa citra kartun dengan resolusi 960 x 540. Hasil pengujian dapat dilihat pada gambar 7 sampai gambar 10, dan tabel 2. Pengujiann berupa jumlah titik yang terdeteksi, yang sesuai dan waktu pencocokan yang dibutuhkan untuk kedua algoritma tersebut.



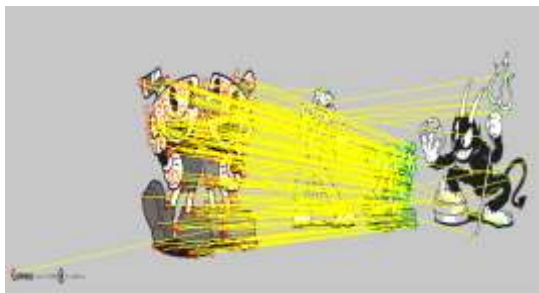
Gbr 7 Fitur yang terdeteksi oleh SIFT



Gbr 8 Fitur yang cocok oleh SIFT



Gbr 9 Fitur yang terdeteksi oleh SURF



Gbr 10 Fitur yang cocok oleh SURF

Gambar 4 menunjukkan hasil fitur yang cocok antara *train image* dan *test image* menggunakan algoritma SIFT. Pencocokan tersebut diindikasikan dengan garis lurus. Sedangkan Gambar 6 merupakan pencocokan menggunakan algoritma SURF.

TABEL II
Hasil Pengujian *Dataset Kedua*

Deskriptor Fitur	Titik Fitur Terdeteksi		Matching Titik Fitur	Waktu yang dibutuhkan untuk Matching Fitur
	Training image	Test image		
SIFT	767	1152	253	14.19 s
SURF	1157	1477	173	3.61

C. Perbandingan Metode SIFT dan SURF

Tabel 3 Perbandingan Metode Algoritma SIFT dan SURF

	SIFT	SURF
Deteksi <i>Keypoint</i>	Perbedaan skala citra dikonvolusi menggunakan fungsi Gaussian	Citra input dikonvolusi menggunakan <i>box filter</i> pada beberapa skala

	<i>Keypoint</i> didapat berdasarkan lokal <i>extrema</i> dalam selisih ruang Gaussian dengan skala yang berbeda-beda	<i>Keypoint</i> ditentukan menggunakan Hessian <i>matrix</i>
Deskripsi <i>Keypoint</i>	Amplitudo gradien area persegi dihitung dengan kekuatan gradien maksimum sebagai arah utama	Respon wavelet Haar digunakan untuk menghitung setiap sektor dalam <i>circular area</i>
	Ekstraksi Fitur dengan membagi <i>region</i> 16x16 menjadi <i>subregion</i> 4x4, untuk setiap <i>subregion</i> dibuat gradien histogram.	Ekstraksi Fitur dengan membagi <i>region</i> 20x20 menjadi <i>subregion</i> 4x4 dan respons Haar wavelet dihitung.
Dimensi Deskriptor	128	64

D. Analisa Recall Dengan Variant Rotasi Test Image Dataset Pertama

TABEL 4
Recall Algoritma SIFT

Photo_test.jpeg	Prediksi Matched Keypoints benar	Jumlah Matched Keypoints	Rasio Keypoints benar	Persentase Keypoints benar
Rotasi 0°	300	300	1	100%
Rotasi 90°	285	285	1	100%
Rotasi 180°	300	300	1	100%

TABEL V
Recall Algoritma SURF

Photo_test.jpeg	Prediksi Matched Keypoints benar	Jumlah Matched Keypoints	Rasio Keypoints benar	Persentase Keypoints benar
Rotasi 0°	52	63	0.825	82.5%
Rotasi 90°	50	62	0.806	80.6%
Rotasi 180°	51	62	0.823	82.3%

E. Analisa Recall Dengan Variant Rotasi Test Image Dataset Kedua

TABEL VI
Recall Algoritma SIFT

Cartoon_test.jpeg	Prediksi Matched Keypoints benar	Jumlah Matched Keypoints	Rasio Keypoints benar	Persentase Keypoints benar
Rotasi 0°	250	253	0.988	99.8%
Rotasi 90°	252	256	0.984	98.4%
Rotasi 180°	244	256	0.953	95.3%

TABEL VII
Recall Algoritma SURF

Cartoon_test.jpeg	Prediksi <i>Matched</i> <i>Keypoints</i> benar	Jumlah <i>Matched</i> <i>Keypoints</i>	Rasio <i>Keypoints</i> benar	Persentase <i>Keypoints</i> benar
Rotasi 0°	153	173	0.884	88.4%
Rotasi 90°	148	169	0.876	87.6%
Rotasi 180°	146	163	0.896	89.6%

V. KESIMPULAN

Berdasarkan hasil pengujian dan Analisa diperoleh kesimpulan sebagai berikut:

1. Jumlah fitur yang terdeteksi oleh SIFT relatif lebih banyak dengan rata-rata presentase data benar sebesar 100% pada *dataset* pertama dan 97.83% pada *dataset* kedua.
2. Jumlah fitur yang dideteksi oleh SURF sebesar 81.8% untuk *dataset* pertama dan 88.53% untuk *dataset* kedua.
3. Kecepatan komputasi SURF relatif lebih cepat daripada SIFT. Diketahui kecepatan rata-rata algoritma SIFT dalam mencari *matched feature* sebesar 72.545 detik pada *range* resolusi dari 756x1008 (*dataset* pertama) sampai 960x540 (*dataset* kedua), dan SURF sebesar 6.58 detik.
4. Dalam hal akurasi SIFT lebih baik daripada SURF, namun dalam hal kecepatan komputasi SURF lebih baik daripada SIFT

REFERENSI

- [1] Kadir, Abdul & Adhi Susanto. (2013). **Teori Dan Aplikasi Pengolahan Citra**. Yogyakarta: Penerbit ANDI
- [2] David G Lowe, **Distinctive Image Features from Scale-Invariant Keypoints**. International Journal of Computer Vision, 2004.
- [3] Bay H, Tuytelaars T, Gool LV. **Speeded-Up Robust Features (SURF)**. Computer Vision and Image Understanding. 2008.