



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk mengubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

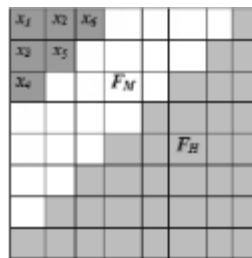
This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

BAB II

TINJAUAN PUSTAKA

2.1.1 Discrete Cosine Transform (DCT)

Pada tahun 2016, telah dibahas pengenalan citra wajah dengan penerapan algoritma DCT (A.Thamizharasi and Jayasudha, 2016). Dengan merepresentasikan sebuah citra ke dalam frekuensi domain, dapat dilakukan suatu teknik pengolahan untuk meningkatkan kualitas citra. DCT merupakan suatu metode yang bekerja berdasarkan frekuensi domain dari sebuah citra. Pada frekuensi domain, pita frekuensi tinggi merepresentasikan *derau* yang ada pada citra. Dengan menghilangkan *derau*, fitur-fitur pada suatu citra dapat dipulihkan. Domain frekuensi dari sebuah citra dapat dibagi menjadi 8x8 blok dan direpresentasikan ke dalam 3 pita frekuensi: pita frekuensi rendah, pita frekuensi sedang, dan pita frekuensi tinggi. Detail area frekuensi tersebut dapat dilihat pada Gambar 2.1 x_1-x_6 menyimbolkan frekuensi rendah, FM melambangkan frekuensi sedang, dan FH melambangkan frekuensi tinggi.



Gambar 2. 1 Frekuensi Domain (Saboori and Hosseini, 2014)

Perhitungan yang digunakan untuk pengaplikasian DCT pada suatu citra didefinisikan sebagai berikut:

$$Z(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) * \cos \left[\frac{(2x+1)u\pi}{2M} \right] \cos \left[\frac{(2y+1)v\pi}{2N} \right] \quad (2. 1)$$

dimana,

$$\alpha(u) = \begin{cases} \frac{1}{\sqrt{M}}, u=0 \\ \sqrt{\frac{2}{M}}, 1 \leq u \leq M-1 \end{cases}$$

dan (2.2)

$$\alpha(v) = \begin{cases} \frac{1}{\sqrt{N}}, v=0 \\ \sqrt{\frac{2}{N}}, 1 \leq v \leq N-1 \end{cases}$$

Nilai konstanta yang terletak pada bagian kiri atas dikenal sebagai basis fungsi DC. DC berperan dalam menentukan keseluruhan pencahayaan yang ada pada citra (Chen, Er and Wu, 2006). Oleh karena itu, keseragaman nilai pada DC dapat diatur dengan rumus berikut:

$$C(0,0) = \log \mu * \sqrt{MN} \quad (2.3)$$

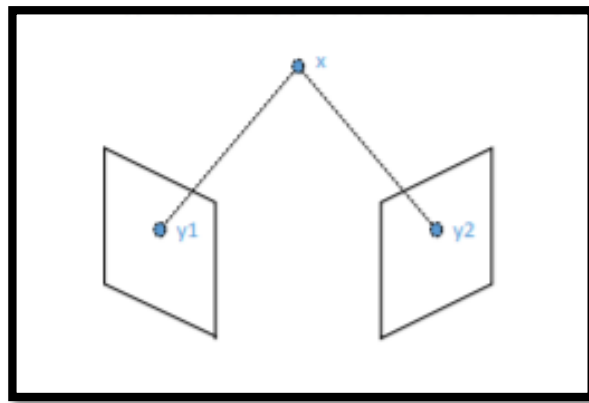
Inverse Discrete Cosine Transform didefinisikan sebagai berikut:

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v)Z(u, v) * \cos \left[\frac{(2x+1)u\pi}{2M} \right] \cos \left[\frac{(2y+1)v\pi}{2N} \right] \quad (2.4)$$

Dengan M dan N adalah ukuran matrix DCT 2D MxN. f(x,y) merepresentasikan koefisien data blok yang asli sedangkan Z(u,v) merepresentasikan koefisien dalam blok setelah menerapkan DCT. DCT diimplementasikan pada suatu citra dengan rumus tersebut sehingga citra yang tadinya dalam frekuensi spasial berubah menjadi frekuensi domain. Pada frekuensi domain terdapat *derau* yang dapat dihilangkan dengan merubah nilai pixel *derau* menjadi 0. Citra wajah yang sudah diubah menjadi frekuensi domain dikembalikan menjadi frekuensi spasial dengan melakukan *inverse*. Melalui DCT, korelasi antara setiap blok dapat diekstraksi dengan lebih baik. Dengan demikian DCT dianggap sebagai *quasi-optimal transform*.

2.2 Epipolar Geometry

Epipolar geometry dapat menghubungkan 2 objek dari adegan yang sama. Metode ini dapat membantu informasi geometris yang ada pada dua citra. Dalam penelitian ini, *Epipolar* efektif dalam membantu pengenalan wajah. Penelitian Carlos juga telah menyarankan pengenalan wajah menggunakan pencocokan stereo untuk menilai kesamaan pada dua titik dalam citra 2D (Castillo and Jacobs, 2009).



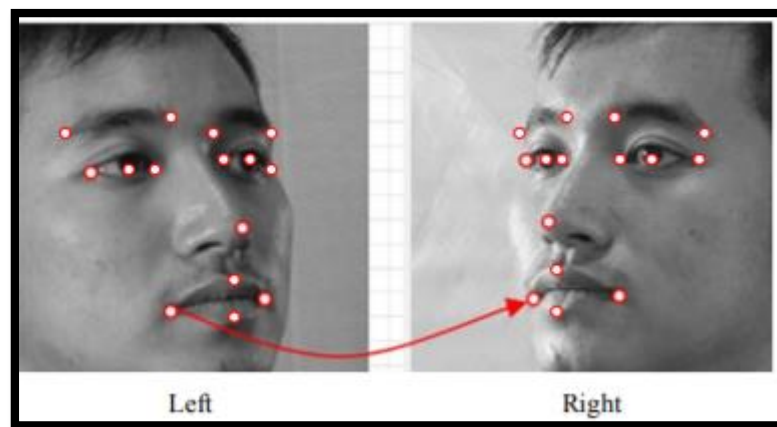
Gambar 2. 2 Epipolar Geometri (Kusnadi *et al.*, 2018)

Titik x (titik 3D) dalam ruang 3D diproyeksikan ke bidang gambar masing-masing, menghasilkan dua gambar yang sesuai poin y_1 dan y_2 . Gambar 2.2 adalah kasus *epipolar geometry* yang ideal. Dalam praktiknya, titik gambar y_1 dan y_2 tidak bisa diukur dengan akurasi sewenang-wenang. Meskipun demikian, itu masi perlu titik yang sama untuk rekonstruksi 3D (Kusnadi *et al.*, 2018).

2.3 Keypoints

Keypoints dapat digunakan untuk rekonstruksi model 3D. Banyak penelitian sebelumnya juga telah membahas tentang titik kunci. Yue Wang dan Yang Song telah membahas *facial keypoints* (Song, 2014). Data yang digunakan relatif besar, yaitu 7049, yang terdiri dari matriks citra dan 15 titik wajah utama. *Keypoints* adalah masalah yang sangat menantang. Fitur wajah sangat bervariasi dari satu orang ke orang lain, dan bahkan untuk satu orang, ada banyak variasi karena pose

3D, ukuran, posisi, sudut pandang, kondisi pencahayaan, kontras, dan bahkan faktor psikologis seperti emosi. Berdasarkan penelitian sebelumnya (Kusnadi *et al.*, 2018), *keypoints* yang digunakan adalah 15 yang merujuk pada suatu peron bernama Kaggle yang tentunya data yang digunakan sah dan telah diteliti oleh banyak orang yang berkecimpung di dunia ilmiah. Citra wajah yang digunakan terdiri dari dua sisi. Namun, dalam penelitian Yue Wang dan Yang Song hanya berkonsentrasi pada posisi pusat mata kiri dan kanan. Dalam penelitian ini lima belas poin utama akan terdeteksi. Berikut adalah contoh *keypoints* pada citra wajah.



Gambar 2. 3 Corresponding points (Kusnadi *et al.*, 2018)

Tabel 2. 1 15 Keypoints

<i>Left eye center</i>	<i>Right eye center</i>
<i>Left eye inner corner</i>	<i>Right eye inner corner</i>
<i>Left eye outer corner</i>	<i>Right eye outer corner</i>
<i>Left eyebrow inner end</i>	<i>Right eyebrow inner end</i>
<i>Left eyebrow outer end</i>	<i>Right eyebrow outer end</i>
<i>Mouth left corner</i>	<i>Mouth right corner</i>
<i>Mouth center bottom and Nose lip</i>	<i>Mouth center top lip</i>

2.4 Detektor Fitur untuk Citra Wajah

Pada penelitian sebelumnya telah diusulkan lima detektor fitur (Kusnadi *et al.*, 2018), yaitu *Speeded Up Robust Features* (SURF), *Features From Accelerated Segment Test* (FAST), *Binary Robust Invariant Scalable Keypoints* (BRISK), *Harris-Stephans*, dan *Minimum Eigen Value*, dan. Dalam mengaplikasikan detektor fitur, yang menjadi acuan pengukuran adalah *precision* dan *recall*. Pada penelitian yang telah dilakukan (Kusnadi *et al.*, 2018), nilai *recall* dari detektor fitur masih rendah. Di sisi lain, terdapat detektor fitur lainnya yang memiliki performa yang cukup baik. Dalam hal ini terlihat pada detektor *Harris-Stephans* karena memiliki nilai *f-score* tertinggi sebesar 0.46. Kelima fitur ini digunakan untuk mendeteksi fitur-fitur wajah dan digunakan satu persatu pada setiap citra yang sama sehingga dapat dilihat hasil yang terbaik dari kelima detektor. Berikut adalah penjelasan lebih lanjut mengenai lima detektor fitur:

2.4.1 SURF

SURF menggunakan pendekatan matriks Hessian yang beroperasi pada gambar integral untuk menemukan titik bunga, yang mengurangi waktu perhitungan secara drastis. SURF adalah fitur invarian rotasi skala yang berisi detektor titik bunga dan deskriptor. Detektor menempatkan titik-titik minat pada citra, kemudian mendeskripsikan fitur titik-titik minat dan menyusun vektor fitur titik-titik bunga. Deskriptor SURF 64-dimensi berfokus pada distribusi spasial informasi gradien di dalam titik penting di sekitar lokasi. Deskriptor SURF tidak berbeda dengan rotasi, skala, kecerahan dan kontras.

Detektor SURF memfokuskan perhatiannya pada struktur seperti gumpalan pada gambar. Struktur-struktur ini dapat ditemukan di sudut-sudut objek, tetapi juga

di lokasi-lokasi dimana refleksi cahaya pada permukaan *specular* lebih kecil. Gambar integral digunakan dalam pendekatan matriks Hessian, yang mengurangi waktu komputasi secara drastis. Diberi titik $X = (x,y)$ di dalam gambar I ,matriks *hessian* $H(x,\sigma)$ dalam X pada skala σ didefinisikan sebagai berikut:

$$H(X, \sigma) = \begin{bmatrix} L_{xx}(X,\sigma) & L_{xy}(X,\sigma) \\ L_{xy}(X,\sigma) & L_{yy}(X,\sigma) \end{bmatrix} \quad (2.5)$$

$L_{xx}(X, \sigma)$, $L_{xy}(X, \sigma)$, dan $L_{yy}(X, \sigma)$ adalah konvolusi dari turunan *Gaussian* urutan kedua $\delta^2 g(\sigma)/\delta x^2$ dengan citra pada titik $X = (x, y)$. Derivatif ini dikenal sebagai *Laplacian of Gaussians*. Maksimal lokal dari respons filter ini terjadi di L_{xx} dan L_{yy} yang sangat positif, dan di L_{xy} yang sangat negatif. Oleh karena itu, ekstrem ini terjadi di dalam citra dengan variasi intensitas tinggi dalam berbagai arah, serta pada titik sadel. Secara visual, berarti bahwa struktur seperti gumpalan mengacu pada sudut dan bintik-bintik. Semua perhitungan turunan ini ada pada skala yang berbeda secara komputasi(Carro *et al.*, 2015). Tujuan dari algoritma SURF adalah membuat proses ini se-efisien mungkin. *SURF* juga memiliki kecepatan yang baik dalam pencocokan dan sangat cocok untuk pengenalan wajah(Du, Su and Cai, 2009).

2.4.2 FAST

FAST digunakan untuk mengidentifikasi tempat menarik di suatu citra. Poin minat ini memiliki info kuat tentang fitur lokal. *FAST* juga dapat memberikan kinerja deteksi yang lebih efektif. Oleh karena itu, detektor *FAST* juga dipilih untuk digunakan dalam penelitian ini. FAST bekerja melalui pendeteksian titik minat dengan mengambil satu pixel pada satu waktu dan mengevaluasi semua pixel dalam lingkaran. Titik *pixel* yang dipilih menyimpan *pixel* lainnya sebagai *vector*. Setiap titik akan dipilih secara berulang sampai semua titik pernah dipilih. Pada detektor

fitur FAST digunakan algoritma ID3 yang bekerja berdasarkan prinsip minimalisasi entropi. Pada semua *pixel* citra akan dikueri sedemikian rupa hingga ditemukan bunga poin dengan jumlah minimum atau yang memiliki informasi paling banyak tentang *pixel*. Entropi untuk himpunan titik yang dipilih dapat secara matematis direpresentasikan sebagai berikut:

$$H(p) = (c + \bar{c}) - c \log_2 c \quad (2.6)$$

Dimana

$$c = |\{p|K_p \text{ adalah benar}\}| \text{ (jumlah sudut)} \quad (2.7)$$

Dan

$$\bar{c} = |\{p|K_p \text{ adalah benar}\}| \text{ (jumlah bukan sudut)} \quad (2.8)$$

H merupakan entropi untuk titik-titik yang dipilih sedangkan p adalah titik yang sedang dipilih untuk dapat dievaluasi. Selain itu, c yang ada adalah jumlah sudut yang didapat pada suatu citra (Viswanathan, 2011).

2.4.3 BRISK

BRISK terdiri dari pendeteksian fitur dan pencocokan *keypoints*. Modularitas BRISK dapat memberikan fleksibilitas dengan deskriptor utama lainnya yang membuat kinerja menjadi optimal. Berfokus pada perhitungan yang efisien, BRISK juga dapat bekerja dalam tingkat kecepatan yang baik. Konsep dari BRISK adalah memanfaatkan pola yang digunakan untuk percobaan ketetangaan dari titik kunci. Penempatan pola disesuaikan untuk titik kunci dalam citra yang didefinisikan dengan rumus berikut:

$$\frac{N(N-1)}{2} \quad (2.9)$$

N merupakan jumlah point *uniform distribution*. Nilai intensitas yang digunakan pada titik ini untuk memperkirakan gradien lokal $g(p_i, p_j)$ adalah $I(p_i, \sigma_i)$ dan $I(p_j, \sigma_j)$ dengan rumus:

$$g(p_i, p_j) = (p_i - p_j) \cdot \frac{I(p_j, \sigma_j) - I(p_i, \sigma_i)}{\|p_i - p_j\|} \quad (2.10)$$

Dalam mengaplikasikan pola percobaan maka pola dirotasi dengan orientasi α dari titik kunci dimana g_y, g_x merupakan gradien pada titik y dan x

$$\alpha = \text{atan2}(g_y, g_x) \quad (2.11)$$

Kemudian bentuk deskriptor biner dengan pencocokan jarak terdekat. Setiap bit dalam deskriptor biner ditentukan dari pencocokan jarak dekat (Leutenegger Stefan, Margarita Chli, 2011).

2.4.4 Harris-Stephans

Harris-Stephans adalah metode poin bunga yang diperkenalkan oleh Moravec. Dalam detektor ini, ada titik menarik yang merupakan lingkungan citra dimana teksturnya berubah secara signifikan ke segala arah (Faille Thesis, 2006). Persamaan algoritma yang dipakai *Harris-Stephans* dapat didefinisikan sebagai berikut:

$$M = G(S) \theta \begin{pmatrix} g_x^2 & g_x g_y \\ g_x g_y & g_y^2 \end{pmatrix} \quad (2.12)$$

g_x adalah gradien dalam arah x; g_y adalah gradien dalam arah y; $G(S)$ merepresentasikan *Gaussian*. Fungsi respon sudut pada algoritma *Harris* adalah sebagai berikut:

$$R = \det M - k \text{tr}^2 M \quad (2.13)$$

R adalah fungsi respons sudut yang diperlukan; $\det M$ adalah penentu Matrix; $\text{tr} M$ adalah jejak Matrix; k adalah konstanta standar, dan umumnya 0.04-

0.06. Dalam praktiknya, nilai tengah R dari suatu gambar dihitung dan jika nilainya maksimum dalam lingkungan dan lebih besar dari ambang batas yang diberikan, titik tersebut dianggap sebagai titik sudut (Wang *et al.*, 2012).

2.4.5 Minimum Eigen Value

Detektor fitur *Minimum Eigen Value* merupakan pengembangan dari detektor fitur *Harris-Stephans*. *Minimum Eigen Value* diperkenalkan oleh Shi dan Carlo Tomasi oleh karena itu detektor ini juga akrab dikenal dengan nama Shi-Tomasi. Perbedaan antara detektor *Harris-Stephans* dan *Minimum Eigen Value* dapat dilihat dari kriteria dalam memilih sudut. *Harris-Stephans* menghitung semua *pixel* kemudian membandingkan skor dengan nilai tertentu. Nilai R dihitung oleh 2 nilai *eigen* sedangkan pada detektor fitur *Minimum Eigen Value* mengusulkan untuk membuang beberapa fungsi pada detektor *Harris-Stephans* dan hanya menggunakan nilai *eigen* untuk menguji pixel. Shi-Tomasi membuktikan secara empiris bahwa skor yang dihasilkan lebih baik (Kadhim and Araheemah, 2019). R didapatkan dari nilai minimal antara 2 nilai *eigen* yang didefinisikan sebagai λ_1, λ_2 .

$$R = \min(\lambda_1, \lambda_2) \quad (2.14)$$

2.5 Grayscale

Citra *grayscale* sangat sering digunakan dalam pengenalan wajah. Penelitian sebelumnya telah banyak membahas tentang skala abu-abu. Juwei Lu mengusulkan untuk menggunakan skala abu-abu dalam pengenalan wajah (Lu and Plataniotis, 2009). Citra *grayscale* digunakan karena sebagian besar perangkat keras pemotretan hanya mendukung citra 8-bit. Selain itu, citra skala abu-abu cukup untuk mengenali wajah, sehingga tidak perlu warna kompleks lainnya (Macedo, Melo and Kelner, 2015).

2.6 Evaluasi Kriteria

Untuk mengevaluasi kinerja detektor fitur, skor pengulangan untuk pasangan gambar tertentu dihitung. Hal tersebut didasarkan pada jumlah kecocokan yang benar dan salah yang diperoleh untuk pasangan gambar. Untuk melakukannya dapat menggunakan *recall* dan *precision*. Recall adalah jumlah daerah yang cocok dengan benar dan jumlah wilayah yang sesuai antara dua gambar yang di tempat kejadian yang sama (Kusnadi *et al.*, 2018). Variabel *True Positives* (*tp*) didefinisikan sebagai *correct match* atau *keypoints* yang sama antara sepasang citra sedangkan *False Negatives* (*fn*) didefinisikan sebagai titik selain *True Positives* yang terdeteksi oleh sistem (Manning, Schütze and Weikurn, 2002).

$$recall = \frac{tp}{tp+fn} \quad (2.15)$$

Precision adalah jumlah pencocokan yang relatif benar terhadap total dari semua kesamaan dapat direpresentasikan sebagai:

$$precision = \frac{tp}{tp+fp} \quad (2.16)$$

False Positives (*fp*) didefinisikan sebagai titik selain *keypoints* yang cocok antara sepasang citra atau lebih dikenal sebagai *correspondences*.

F-Score adalah ukuran yang menggabungkan *precision* dan *recall* dengan nilai rata-rata yang harmonis. Rumusnya adalah:

$$F - Score = 2x \frac{(Recall \times Precision)}{(Recall + Precision)} \quad (2.17)$$

Batas nilai *F-Score* adalah antara 0 dan 1