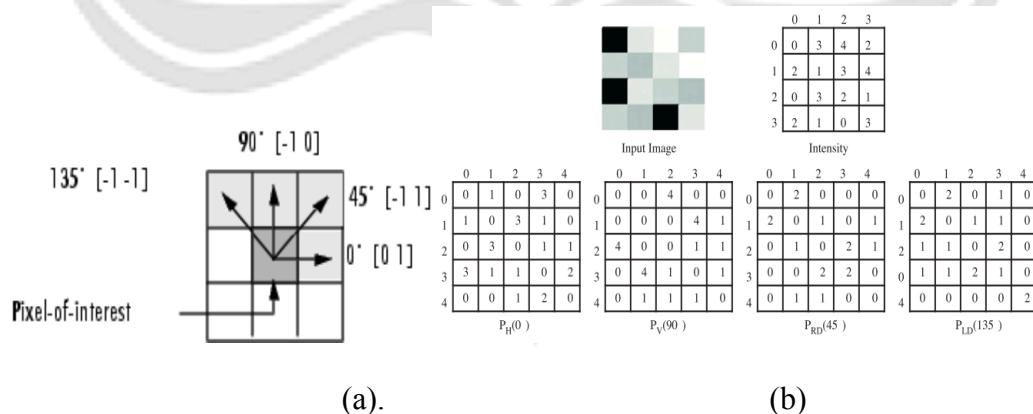


## BAB III

### LANDASAN TEORI

#### 3.1 Metode GLCM ( Gray Level Co-Occurrence Matrix)

Metode GLCM menurut Xie dkk (2010) merupakan suatu metode yang melakukan analisis terhadap suatu piksel pada citra dan mengetahui tingkat keabuan yang sering terjadi. Metode ini juga untuk tabulasi tentang frekuensi kombinasi nilai piksel yang muncul pada suatu citra . Untuk melakukan analisis citra berdasarkan distribusi statistik dari intensitas pikselnya, dapat dilakukan dengan mengekstrak fitur teksturnya (Pullaperuma & Dharmaratne 2013). GLCM merupakan suatu metode untuk melakukan ekstraksi ciri berbasis statistik, perolehan ciri diperoleh dari nilai piksel matrik, yang mempunyai nilai tertentu dan membentuk suatu sudut pola (Kasim & Harjoko., 2014), (Xie dkk. 2010). Untuk sudut yang dibentuk dari nilai piksel citra menggunakan GLCM adalah  $0^0$ ,  $45^0$ ,  $90^0$ ,  $135^0$  (Eleyan & Demirel 2011), untuk sudut yang terbentuk terlihat pada Gambar 3.1.



Gambar 3. 1 (a) Piksel dengan berbagai sudut (b) Ilustrasi Matriks co-ocurensi

Sumber : (Eleyan & Demirel 2011)

Dari piksel-piksel tersebut terbentuk matrik ko-okurensi dengan pasangan pikselnya. Adanya matrik tersebut berdasarkan kondisi bahwa suatu matrik piksel akan mempunyai nilai perulangan sehingga terdapat pasangan aras keabuannya (Thakare & Patil 2014). Kondisi nilai piksel tersebut dinotasikan sebagai matrik dengan jarak dua posisi  $(x_1, y_1)$  dan  $(x_2, y_2)$ . Berdasarkan kondisi tersebut terlihat bahwa untuk membedakan antar matrik gambar dapat dilihat berdasarkan ciri matrik dengan menggunakan persamaan Maurya dkk. (2014) sebagai berikut;

Tabel 3. 1 Formula Ekstraksi Ciri

| No | Feature     | Formula   |
|----|-------------|---|
| 1  | Contras     | $\sum_i \sum_j (i - j)^2 C(i, j)$ (3.1)           |
| 2  | Energi      | $\sum_i \sum_j C^2(i, j)$ (3.2)                   |
| 3  | Entropy     | $\sum_i \sum_j C(i, j) \log(C(i, j))$ (3.3)       |
| 4  | Homogenitas | $\sum_i \sum_j \frac{C(i, j)}{1 +  i - j }$ (3.4) |

sumber : (Lee & Choi 2010)

Berdasarkan Tabel 3.1, untuk *feature* kontras penyebaran elemen matrik pada citra terletak jauh dari diagonal utama akan mempunyai nilai kontras yang cukup besar. Nilai kontras merupakan suatu variasi antar derajat keabuan di suatu matrik pada citra Maurya dkk. (2014). Pada citra juga dapat diketahui nilai energi atau *angular second moment*. Untuk nilai energi tersebut ditunjukkan pada  $C^2$  pada baris dan kolomnya  $(i, j)$ . Sedangkan untuk menghitung nilai entropi dapat diketahui dari ketidakteraturan bentuk, dan pada derajat keabuan merata akan mempunyai nilai besar dan bernilai kecil pada saat piksel citra beragam dengan berbagai varian atau tidak teratur. Untuk suatu matrik citra, nilai homogenitasnya ditentukan dari derajat keabuan yang sejenis.

### 3.2 Deteksi Tepi

Pada proses manipulasi suatu matrik untuk mendapatkan nilai derajat keabuan dari suatu himpunan piksel yang terletak pada boundary (dalam arah vertikal), masing-masing terletak pada transisi *step orthogonal* dari tingkat keabuan. Ketidaksempurnaan optik, sampling, dan proses pengambilan data citra, akan menghasilkan tepi-tepi yang kabur, dengan derajat kekaburan ditentukan oleh faktor-faktor seperti kualitas peralatan yang digunakan untuk mengambil data citra, rata-rata sampling, dan kondisi pencahayaan (Yu dkk. 2012). Akibatnya, tepi lebih banyak dimodelkan seperti “ramp”. Ketebalan tepi ditentukan oleh panjang ramp. Panjang ramp ditentukan oleh kemiringan (*slope*), dan *slope* ditentukan oleh derajat kekaburan. Tepian yang kabur cenderung lebih tebal, dan tepian yang tajam cenderung lebih tipis. Magnitude dari turunan pertama bias digunakan untuk mendeteksi keberadaan tepi pada suatu titik dalam citra (misalnya, menentukan apakah suatu titik berada pada ramp atau tidak). Tanda dari turunan kedua bisa digunakan untuk menentukan apakah suatu piksel tepi terletak pada sisi gelap atau sisi terang dari tepi (Kasim & Harjoko.,2014).

Property *zero-crossing* (garis lurus imajiner yang menghubungkan nilai ekstrim positif dan negatif dari turunan kedua akan melintasi nol di pertengahan tepi) cukup berguna untuk menentukan pusat dari tepi yang tebal. Agar dapat diklasifikasikan sebagai titik tepi, transisi tingkat keabuan pada titik tersebut harus cukup kuat dibandingkan background di sekitarnya. Untuk menentukan apakah suatu nilai “cukup signifikan” atau tidak, bisa digunakan *threshold*. Jadi, suatu

titik di dalam citra merupakan bagian dari tepi, jika turunan pertama 2-D nya lebih besar dari *threshold*. Himpunan titik-titik yang terhubung menurut kriteria keterhubungan tertentu didefinisikan sebagai tepi. Istilah *segmen edge* digunakan jika ukuran tepi relatif pendek dibanding ukuran citra. Permasalahan dalam segmentasi adalah bagaimana cara merangkai segmen-segmen tepi ini menjadi tepi yang lebih panjang. Tepi juga bisa ditentukan menggunakan *property zero crossings* dari turunan kedua. Analisis citra pada dasarnya terdiri dari tiga tahapan, yaitu ekstraksi ciri, segmentasi, dan klasifikasi (Yu, 2010). Faktor kunci dalam mengekstraksi ciri adalah kemampuan mendeteksi keberadaan tepi di dalam citra. Ada beberapa metode deteksi tepi. Penggunaan metode deteksi tepi yang tidak tepat, akan menghasilkan pendeteksian yang gagal. Pendeteksian tepi merupakan tahapan untuk melingkupi informasi di dalam citra. Tepi mencirikan batas objek dan karena itu tepi berguna untuk proses segmentasi dan identifikasi objek di dalam citra.

Untuk langkah deteksi tepi menggunakan metode Canny menurut (Kasim & Harjoko.,2014) dengan tahapan sebagai berikut (1) Memfilter noise pada citra dengan mengimplementasikan Filter Gaussian, dan citra tampak buram. (2) Deteksi tepi dengan salah satu operator deteksi ( contoh Roberts, Perwitt atau Sobel) dan dilakukan pencarian horizontal maupun vertikal. (3) Menentukan arah tepian. (4) Memperkecil garis tepi yang muncul, menerapkan non maximum suppression sehingga menghasilkan garis tepian yang lebih tajam. (5) Binerisasi dengan menerapkan dua buah nilai ambang batas.

### 3.3 Fuzzy C-Mean(FCM)

Fuzzy C-means Clustering (FCM) (Sujit et al. 2012), atau dikenal juga sebagai Fuzzy ISODATA, merupakan salah satu metode clustering yang merupakan bagian dari metode Hard K-Means. FCM menggunakan model pengelompokan fuzzy sehingga data dapat menjadi anggota dari semua kelas atau cluster terbentuk dengan derajat atau tingkat keanggotaan yang berbeda antara 0 hingga 1. Tingkat keberadaan data dalam suatu kelas atau cluster ditentukan oleh derajat keanggotaannya. Teknik ini pertama kali diperkenalkan oleh Jim Bezdek pada tahun 1981. Konsep dasar FCM, pertama kali adalah menentukan pusat cluster yang akan menandai lokasi rata-rata untuk tiap-tiap cluster. Pada kondisi awal, pusat cluster ini masih belum akurat. Tiap-tiap data memiliki derajat keanggotaan untuk tiap-tiap cluster. Dengan cara memperbaiki pusat cluster dan nilai keanggotaan tiap-tiap data secara berulang, maka dapat dilihat bahwa pusat cluster akan menuju lokasi yang tepat. Fungsi obyektif yang digunakan FCM adalah

$$J(U, V; X) = \sum_{k=1}^n \sum_{i=1}^C (\mu_{ik})^w (d_{ik})^2 \quad (3.5)$$

dengan  $w \in [1, \infty]$ ,

$$d_{ik} = d(x_k - v_i) = \left[ \sum_{j=1}^m (x_{kj} - v_{ij})^2 \right]^{1/2} \quad (3.6)$$

X adalah data yang dicluster :

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix} \quad (3.7)$$

dan V adalah matriks pusat cluster :

$$V = \begin{bmatrix} v_{11} & \cdots & v_{1m} \\ \vdots & & \vdots \\ v_{c1} & \cdots & v_{cm} \end{bmatrix} \quad (3.8)$$

Nilai  $J_w$  terkecil adalah yang terbaik, sehingga :

$$J_w^*(U^*, V^*; X) = \min_{M_{jk}} J(U, V, X) \quad (3.9)$$

Algoritma FCM secara lengkap diberikan sebagai berikut (Takeda et al. 2008) :

Tentukan :

- Matriks  $X$  berukuran  $n \times m$ , dengan  $n$  = jumlah data yang akan di cluster; dan  $m$  = jumlah variabel (kriteria).
- Jumlah cluster yang akan dibentuk ( $C \geq 2$ )
- Pangkat (pembobot  $w > 1$ )
- Maksimum iterasi
- Kriteria penghentian ( $\varepsilon$  = nilai positif yang sangat kecil)
- Bentuk matriks partisi awal  $U$  (derajat keanggotaan dalam cluster); matriks partisi awal biasanya dibuat secara acak

$$U = \begin{bmatrix} \mu_{11}(x_1) & \mu_{12}(x_2) & \cdots & \mu_{1n}(x_n) \\ \mu_{21}(x_1) & \mu_{22}(x_2) & \cdots & \mu_{2n}(x_n) \\ \vdots & & & \vdots \\ \mu_{c1}(x_1) & \mu_{c2}(x_2) & \cdots & \mu_{cn}(x_n) \end{bmatrix} \quad (3.10)$$

Hitung pusat cluster  $V$  untuk setiap cluster

$$V_{ij} = \frac{\sum_{k=1}^n (\mu_{ik})^w \cdot x_{kj}}{\sum_{k=1}^n (\mu_{ik})^w} \quad (3.11)$$

Perbaiki derajat keanggotaan setiap data pada setiap cluster (perbaiki matriks partisi)

$$\mu c_i(x_i) = \frac{1}{\sum_{j=1}^2 \left( \frac{\|x_1 - v_1\|}{\|x_1 - v_j\|} \right)^2} \quad (3.12)$$

dengan :

$$d_{ik} = d(x_k - v_i) = \left[ \sum_{j=1}^m (x_{kj} - v_{ij}) \right]^{1/2} \quad (3.13)$$

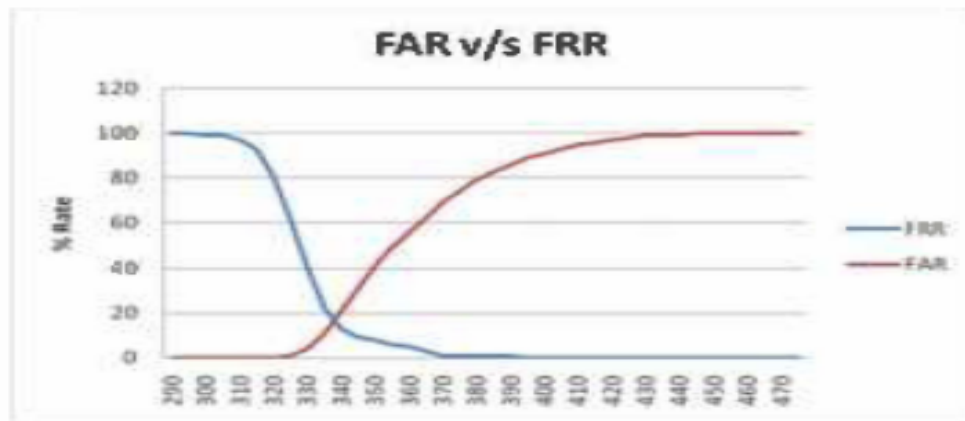
Tentukan kriteria penghentian iterasi, yaitu perubahan matriks partisi pada iterasi sekarang dan iterasi sebelumnya

$$\Delta = \|u^t - u^{t-1}\| \quad (3.14)$$

Apabila  $\Delta < \varepsilon$  maka iterasi dihentikan.

### 3.4 *Equal Error Rate (EER)*

Untuk mengukur kualitas suatu metode identifikasi pada printer forensik dapat dilakukan dengan menghitung nilai FRR (*false rejection rate*) dan nilai FAR (*false acceptance rate*). FRR dapat diartikan sebagai kesalahan dari hasil identifikasi dengan menolak atau menyatakan sampel bukti tidak sama meskipun pada kenyataannya kedua sampel tersebut berasal dari sumber printer yang sama. Sedangkan untuk FAR merupakan kesalahan hasil identifikasi yang menyatakan menerima atau menyatakan kedua sampel dokumen cetak tersebut identik meskipun pada kenyataannya sampel dokumen tersebut berasal dari sumber printer yang berbeda. Untuk FRR dan FAR ditunjukkan pada Gambar 3.2. Untuk nilai FRR maupun nilai FAR mempunyai rentang dari 0 sampai dengan 1.



Gambar 3. 2 ilustrasi Nilai FRR dan FAR  
Sumber (Kekre et al. 2012)

Pertemuan nilai FRR dan FAR, pada titik tersebut dinyatakan sebagai nilai EER (*Equal Error Rate*). Nilai EER merupakan ukuran kualitas dari suatu sistem identifikasi. Untuk nilai FRR dapat dihitung menggunakan persamaan (3.15), dari persamaan tersebut dapat dinyatakan bahwa untuk menghitung FRR dengan mengetahui banyaknya hasil penolakan identifikasi terhadap data yang seharusnya dan dibandingkan dengan keseluruhan hasil pengujian.

$$FRR = \frac{\text{Jumlah yang salah tolak}}{\text{Jumlah keseluruhan kejadian}} \times 100\% \quad (3.15)$$

Sedangkan untuk menghitung nilai FAR dengan menggunakan persamaan (3.16), untuk nilai FAR merupakan perhitungan jumlah kesalahan identifikasi yang menyatakan diterima dibandingkan dengan jumlah pengujian yang dilakukan.

$$FAR = \frac{\text{Jumlah kejadian yang salah diterima}}{\text{Jumlah keseluruhan kejadian}} \times 100\% \quad (3.16)$$