

## AKUISISI *FOREGROUND* DAN *BACKGROUND* BERBASIS FITUR DCT PADA *MATting* CITRA SECARA OTOMATIS

Meidya Koeshardianto<sup>\*1</sup>, Eko Mulyanto Yuniarno<sup>2</sup>, Mochamad Hariadi<sup>3</sup>

<sup>1</sup>Jurusan Teknik Informatika, Fakultas Teknik, Universitas Trunojoyo Madura

<sup>2,3</sup>Departemen Teknik Elektro, Fakultas Teknologi Elektro, Institut Teknologi Sepuluh Nopember Surabaya

Email: <sup>1</sup>meidya@trunojoyo.ac.id, <sup>2</sup>ekomulyanto@ee.its.ac.id, <sup>3</sup>mochar@ee.its.ac.id

<sup>\*</sup>Penulis Korespondensi

(Naskah masuk: 10 Juli 2019, diterima untuk diterbitkan: 27 April 2020)

### Abstrak

Teknik pemisahan *foreground* dari *background* pada citra statis merupakan penelitian yang sangat diperlukan dalam *computer vision*. Teknik yang sering digunakan adalah *image segmentation*, namun hasil ekstraksinya masih kurang akurat. *Image matting* menjadi salah satu solusi untuk memperbaiki hasil dari *image segmentation*. Pada metode *supervised*, *image matting* membutuhkan *scribbles* atau *trimap* sebagai *constraint* yang berfungsi untuk melabeli daerah tersebut adalah *foreground* atau *background*. Pada makalah ini dibangun metode *unsupervised* dengan mengakuisisi *foreground* dan *background* sebagai *constraint* secara otomatis. Akuisisi *background* ditentukan dari varian nilai fitur DCT (*Discrete Cosinus Transform*) yang dikelompokkan menggunakan algoritme *k-means*. Untuk mengakuisisi *foreground* ditentukan dari subset hasil kluster fitur DCT dengan fitur *edge detection*. Hasil dari proses akuisisi *foreground* dan *background* tersebut dijadikan sebagai *constraint*. Perbedaan hasil dari penelitian diukur menggunakan MAE (*Mean Absolute Error*) dibandingkan dengan metode *supervised matting* maupun dengan metode *unsupervised matting* lainnya. Skor MAE dari hasil eksperimen menunjukkan bahwa nilai *alpha matte* yang dihasilkan mempunyai perbedaan 0,0336 serta selisih waktu proses 0,4 detik dibandingkan metode *supervised matting*. Seluruh data citra berasal dari citra yang telah digunakan para peneliti sebelumnya.

**Kata kunci:** *Image segmenterasi, image matting, fitur DCT, deteksi tepi, scribbles*

## FOREGROUND AND BACKGROUND ACQUISITION FOR AUTOMATIC IMAGE MATTING BASED ON DCT FEATURE

### Abstract

The technique of separating the foreground and the background from a still image is widely used in computer vision. Current research in this technique is image segmentation. However, the result of its extraction is considered inaccurate. Furthermore, image matting is one solution to improve the effect of image segmentation. Mostly, the matting process used scribbles or trimap as a constraint, which is done manually as called a supervised method. The contribution offered in this paper lies in the acquisition of foreground and background that will be used to build constraints automatically. Background acquisition is determined from the variant value of the DCT feature that is clustered using the k-means algorithm. Foreground acquisition is determined by a subset resulting from clustering DCT values with edge detection features. The results of the two stages will be used as an automatic constraint method. The success of the proposed method, the constraint will be used in the supervised matting method. The difference in results from In the research experiment was measured using MAE (Mean Absolute Error) compared with the supervised matting method and with other unsupervised matting methods. The MAE score from the experimental results shows that the alpha matte value produced has a difference of 0.336, and the difference in processing time is 0.4 seconds compared to the supervised matting method. All image data comes from images that have been used by previous researchers.

**Keywords:** *image segmentation, image matting, DCT feature, Laplacian matting*

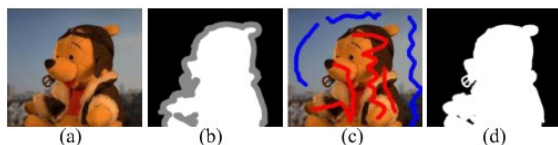
### 1. PENDAHULUAN

Proses pemisahan obyek *foreground* dan *background* pada sebuah citra merupakan teknik

yang sangat diperlukan dalam bidang multimedia khususnya pembuatan visual efek film animasi. *Image segmentation* merupakan salah satu teknik

yang sering digunakan. Namun, ekstraksi *foreground* dan *background* yang dihasilkannya belum dapat memisahkan obyek *foreground* secara akurat. Beberapa pendekatan metode ekstraksi dilakukan salah satunya menggunakan *image matting*, diperkenalkan sebagai solusi untuk memperbaiki hasil ekstraksi yang dilakukan pada *image segmentation*.

Perkembangan *image matting* diawali dengan penggunaan *trimap* dan *scribbles* atau dikenal sebagai *constraint* dengan mengasumsikan warna hijau atau biru sebagai *background* pada sebuah citra. Teknik ini biasa dikenal dengan teknik *green screen matting* atau *blue screen matting*. Namun, beberapa permasalahan ditemui saat penggunaan teknik *green screen matting* diimplementasikan. Permasalahannya adalah dibutuhkan layar yang relatif besar serta sistem pencahayaan yang ideal. Untuk mengatasi permasalahan tersebut diperkenalkan *Natural image matting*.



Gambar 1. (a) Citra asli. (b) *trimap*. (c) *scribbles*. (d) *alpha matte*. (Zhu et al., 2015)

Pendekatan *natural image matting* dilakukan dengan dua cara, yakni: terbimbing (*supervised*) dimana membutuhkan *constraint* dan tidak terbimbing (*unsupervised*) atau secara otomatis. Pada Gambar 1(b), *constraint* dibentuk dengan melabeli hampir seluruh daerah citra dengan nilai 0 sebagai *background* dan nilai 1 sebagai *foreground* atau dikenal dengan *trimap*. Sisanya adalah daerah yang belum diketahui nilainya. Gambar 1(c) adalah *constraint* berupa *scribbles* dimana hanya memberikan sebagian label pada daerah citra. Seluruh *constraint* ditentukan secara manual sehingga mempunyai kelebihan dapat menentukan *foreground* dan *background* secara akurat artinya sesuai dengan asumsi. Namun kelemahannya dibutuhkan usaha lebih untuk melabeli daerah *foreground* dan *background*.

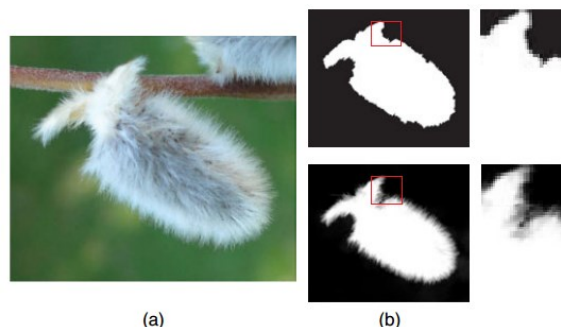
Sangat berbeda dengan *supervised matting*, pada metode *unsupervised* proses *matting* dapat dilakukan tanpa memberikan label pada sebagian daerah *foreground* dan *background*. Dampaknya adalah proses *matting* dapat dilakukan tanpa harus membangun *constraint* sehingga dapat mengurangi kompleksitas pengguna. Meskipun dapat dilakukan secara otomatis, dalam beberapa kasus teknik *supervised* mempunyai kelemahan dalam menghasilkan ekstraksi yang sesuai asumsi pengguna mengenai *foreground* dan *background*.

Pada penelitian kali ini diperkenalkan metode *automatic image matting* dengan mengakuisisi *foreground* dan *background* sebagai *constraint* secara otomatis. Proses pembentukan *constraint*

dilakukan dua tahap yakni akuisisi *background* dan *foreground*. Akuisisi *background* dilakukan dengan jalan mendeteksi besaran varian dari fitur yang dihasilkan oleh *Discrete Cosinus Transform* (DCT). Hasil dari fitur DCT selanjutnya dilakukan proses kluster menggunakan *k-means*. Untuk akuisisi *foreground*, ditentukan dengan metode *edge detection*. Hasil dari akuisisi *foreground* ditentukan oleh daerah kluster *background* yang bukan merupakan subset pada *edge detection*. Hasil kedua proses tersebut digunakan sebagai *constraint* pada *supervised matting*. Dengan demikian, keterbaruan dari penelitian ini adalah membuat metode *supervised matting* menjadi *unsupervised matting* dengan mengakuisisi *foreground* dan *background* sebagai *constraint* secara otomatis.

## 2. PERMASALAHAN IMAGE MATTING

*Image matting* dan *image segmentation* merupakan dua metode yang digunakan untuk memisahkan obyek *foreground* dari citra. Ada perbedaan yang sangat mendasar antara kedua metode ini yakni *alpha matte* yang dihasilkan dari kedua metode tersebut. Untuk *image segmentation*, *alpha matte* yang dihasilkan bernilai 0 dan 1. Berbeda dengan metode *matting*, *alpha matte* yang dihasilkan bernilai antara 0 dan 1. Sehingga ekstraksi yang dihasilkan lebih detail. Perbedaan dapat dilihat pada Gambar.2(b) menampilkan perbedaan *alpha matte* dari metode *image segmentation* dengan *image matting*.



Gambar.2 Perbandingan hasil *alpha matte*. (a) Citra asli. (b, atas) Ekstraksi menggunakan metode segmentasi. (b, bawah) Ekstraksi menggunakan metode *matting* (Wang and Cohen, 2007)

Namun demikian, bukan berarti *image segmentation* telah ditinggalkan oleh para peneliti, dalam beberapa hal *image segmentation* masih digunakan sebagai *preprocessing* di bidang pengenalan wajah dan sidik jari serta identifikasi citra lainnya. Bahkan dalam *image matting*, proses *image segmentation* tetap dapat digunakan untuk membangun *constraint* yakni sebagai *trimap*.

*Image matting* pertama kali diperkenalkan oleh (Porter and Duff, 1984) dengan model persamaan:

$$I_i = \alpha_i F_i + (1 - \alpha_i) B_i \quad (1)$$

dimana sebuah citra  $I$  terletak pada posisi  $i$  merupakan vektor dari masing – masing warna RGB (red, green, blue) yang bercampur sehingga nilai  $\alpha$  relatif proporsional terhadap nilai *foreground*  $F$  dan *background*  $B$ . Nilai  $\alpha$  mendekati 0 maka piksel tersebut dapat dipastikan sebagai *background* demikian juga sebaliknya apabila nilai  $\alpha$  mendekati 1 maka dapat dipastikan bahwa piksel tersebut adalah *foreground*. Hasil akhir dari proses *matting* pada sebuah citra  $I$  adalah citra *foreground*  $F$ , *background*  $B$  dan nilai *alpha matte*  $\alpha$ .

Diawal makalah telah dijelaskan, untuk mempermudah proses *matting* digunakanlah *blue screen matting* (Smith and Blinn, 1996). Hal ini bertujuan untuk mengurangi variabel yang harus dicari. Persamaan (1) menghasilkan tujuh variabel yang harus diketahui dalam menyelesaikan proses *matting* yakni  $\alpha, F^c$  dan  $B^c$  dimana  $c \in \{R, G, B\}$ . Dengan menggunakan metode *blue screen matting*, maka komponen warna biru pada citra dapat dieliminasi sehingga berdasarkan persamaan (1) tersisa tiga variabel yang belum diketahui yakni  $\alpha, F^R$  dan  $F^G$ . Penyelesaian sisa variabel yang belum diketahui dapat diselesaikan menjadi:

$$\alpha = I^B - B^B, I^R = \alpha F^R, I^G = \alpha F^G \quad (2)$$

Persamaan (2) atau metode *blue screen matting* masih digunakan sebagai pembentuk citra *ground truth* pada *image matting*.

Untuk saat ini, penelitian *image matting* dititikberatkan pada metode *natural image matting*. Pada metode *natural image matting* telah berkembang menjadi dua yakni *supervised* dan *unsupervised*.

Dalam metode *supervised*, *scribbles* digunakan pada metode *affinity based matting* (Levin, Lischinski and Weiss, 2008) untuk menyelesaikan *alpha matte*. Variasi *affinity* didefinisikan sebagai nilai bobot pembentuk *alpha matte*. Penggunaan metode *poisson* (Sun *et al.*, 2004) digunakan mengestimasi nilai *alpha matte*. Untuk mengurangi kompleksitas dilakukan modifikasi gradien secara otomatis dan interaktif berdasarkan intensitas (Grady *et al.*, 2005) perubahan baik pada *foreground* dan *background*.

Perkembangan metode *supervised* selanjutnya didekati dengan penggunaan *constraint* secara otomatis menggunakan *trimap* (Vicas *et al.*, 2013; Shen *et al.*, 2016; Cho *et al.*, 2017). *Convolutional Neural Network* (CNN) (Shen *et al.*, 2016) digunakan untuk mendefinisikan komponen *constraint* pada *matting*. Dalam (Cho *et al.*, 2017), *image matting* digunakan sebagai pengenalan obyek dengan memodifikasi klaster geometrik untuk menentukan *foreground* dan *background* sebagai *trimap*. Teknik segmentasi digunakan sebagai awal dari pembentukan *trimap*. Kontur hasil dari *image segmentation* diasumsikan sebagai inisialisasi. Selanjutnya struktur *foreground* dianalisa sehingga

terbentuk *constraint* secara otomatis. Pembentukan *constraint* juga dilakukan menggunakan teknik registrasi pada citra menggunakan SIFT (Koeshardianto, 2016) untuk membangun *constraint* secara otomatis.

Metode *unsupervised matting* pertama kali diperkenalkan (Levin, Rav-acha and Lischinski, 2008) dengan metode *spectral matting*. Akuisisi *background* dan *foreground* ditentukan melalui besaran nilai *eigen* kemudian diimplementasikan pada video (Eisemann, Wolf and Magnor, 2009). Ekstraksi obyek dengan jalan membentuk citra dalam dua atau lebih segmen dimana setiap segmen adalah kumpulan dari varian piksel yang homogen atau dikenal dengan layer. Dari pembentukan layer tersebut beberapa metode menggunakannya sebagai dasar melakukan *soft segmentation* menggunakan global optimisasi (Tan *et al.*, 2013), ekstraksi warna per piksel (Aksoy *et al.*, 2018).

Dengan berkembangnya metode *unsupervised matting*, terjadi pergeseran topik penelitian yang dihadapi khususnya di bidang animasi film, teknik *Rotoscoping* juga melakukan hal yang sama yakni memisahkan *foreground* dan *background* secara interaktif (Lu *et al.*, 2016). *Rotoscoping* menggunakan teknik *Roto++* (Li *et al.*, 2016) merupakan proses *soft segmentation* obyek *foreground* dan *background* dalam video. Untuk mempermudah proses tersebut, dikembangkan teknik ROAM (Miksik *et al.*, 2017) yang merupakan pengembangan Grab-Cut (Rother, Kolmogorov and Blake, 2004). Segmentasi menggunakan *constrained* berdasarkan lokasi, skala serta model dari target digunakan (Long, Liu and Han, 2017) untuk memisahkan obyek dari video. Menggunakan *video segmentation*, (Mashtalir and Mashtalir, 2016) menterjemahkan *low-level* fitur pada video agar mudah dimengerti oleh pengguna. (Zeng *et al.*, 2017) menggunakan *Fuzzy partition* untuk mengekstrak obyek hasil dari *dynamic background subtraction*.

Penelitian mengenai ekstraksi obyek dari video masih terus berkembang hingga saat ini. *Background subtraction*, *video segmentation*, *rotoscoping* merupakan sebagian teknik untuk memisahkan obyek dari video dan pada penelitian yang berkembang saat ini telah dilakukan secara otomatis. Dengan berbagai permasalahan serta trend penelitian tersebut, dalam makalah ini dibahas tentang *automatic image matting* sehingga dapat diimplementasikan pada video.

### 3. LAPLACIAN MATTING

Berdasarkan persamaan (1), setiap warna pada posisi ke  $i$  terurai menjadi dua layer yakni *foreground* dan *background* secara linear. Dengan menggunakan fungsi linear,  $\alpha$  dapat ditulis menjadi:

$$\alpha_i \approx aI + b, \forall i \in w \quad (3)$$

Dimana  $a = \frac{1}{F-B}$ , dan  $b = -\frac{B}{F-B}$  dan  $w$  adalah blok piksel. Dengan mencari nilai minimum dari *cost function*, maka koefisien  $a, b$  dan  $\alpha$  dapat diestimasi menjadi:

$$J(\alpha, a, b) = \sum_{j \in i} \left( \sum_{i \in w_j} (\alpha_i - a_j - b_j)^2 + \epsilon a_j^2 \right) \quad (4)$$

dimana  $w_j$  adalah blok piksel disekitar  $j$  dan  $\epsilon$  merupakan konstanta sebagai regularisasi pada  $a_j$ .

Selanjutnya, ukuran blok piksel sebesar  $3 \times 3$  digunakan saat implementasi untuk mendeteksi propagasi intensitas antar piksel tetangga. Menggunakan teorema (1) pada (Levin, Lischinski and Weiss, 2008),  $a_j$  dan  $b_j$  dapat dieliminasi dari persamaan (4) untuk mendapatkan sebuah *cost function* yang tergantung hanya pada  $\alpha$  menjadi:

$$J(\alpha) = \alpha^T L \alpha \quad (5)$$

Variabel  $\alpha$  merupakan vektor berukuran  $N \times 1$  dan  $L$  adalah matrik *Laplacian* berukuran  $N \times N$  sehingga nilai pada posisi  $(i, j)$  dapat didefinisikan menjadi:

$$L_{ij} = \sum_{k | (i,j) \in w_k} \left( \delta_{ij} - \frac{1}{|w_k|} \left( 1 + \frac{\epsilon}{|w_k| + \sigma_k^2} (I_i - \mu_k)(I_j - \mu_k) \right) \right) \quad (6)$$

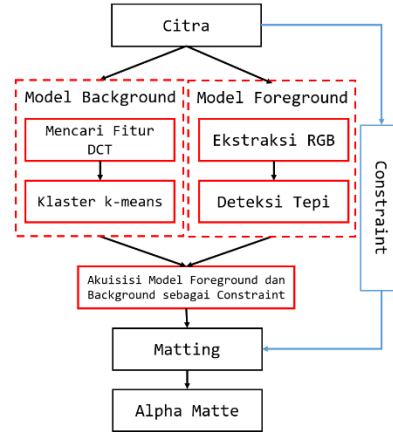
Variabel  $N$  adalah jumlah piksel sebagai obyek target,  $\delta_{ij}$  adalah *kronecker delta*. Sedangkan  $\mu_k$  dan  $\sigma_k^2$  adalah mean dan intensitas varian pada blok piksel  $w_k$ . Untuk  $|w_k|$  adalah jumlah piksel pada blok  $w_k$ . Sehingga *alpha matte* dapat diselesaikan dengan:

$$\alpha = \operatorname{argmin} \alpha^T L \alpha + \lambda (\alpha^T - b_s^T) D_s (\alpha - b_s) \quad (7)$$

Dimana  $\lambda$  adalah konstanta dengan nilai yang besar,  $D_s$  adalah diagonal matrik dengan nilai 1 untuk piksel *constraint* dan 0 untuk piksel yang lainnya. Untuk  $b_s$  merupakan *constraint* dimana terdiri atas nilai spesifik *alpha*. Dalam sistem linear, variabel  $\alpha$  pada persamaan (7) dapat ditulis menjadi:

$$(L + \lambda D_s) \alpha = \lambda b_s \quad (8)$$

Dalam metode *supervised*,  $b_s$  ditentukan secara manual atau dalam *image matting* dikenal dengan *scribbles*. Dalam penelitian ini, persamaan (8) digunakan sebagai formula proses *matting*. Variabel  $b_s$  didekati dengan mengakuisisi *foreground* dan *background* sebagai *scribbles* serta dilakukan secara otomatis.



Gambar. 3 Diagram alir penelitian. Blok warna merah adalah kontribusi penelitian dalam membangun *constraint* secara otomatis

#### 4. METODOLOGI AKUISISI *FOREGROUND* DAN *BACKGROUND*

Akuisisi atau proses mendefinisikan *foreground* dan *background* digunakan untuk membangun *constraint* pada *image matting*. Diagram alir pada Gambar. 3 memperlihatkan tahapan penelitian yang akan dilakukan yakni: membangun model *background* dan model *foreground*. Langkah selanjutnya adalah mengakuisisi dari kedua model yang telah dibangun. Garis warna biru pada Gambar. 3 adalah proses yang dilakukan secara manual yakni *constraint* ditentukan oleh pengguna sebelum proses *matting* dilakukan.

##### 4.1. Model Background

**Ekstraksi Fitur:** DCT dipilih karena pada prosesnya tidak memerlukan memory besar serta penentuan fitur dapat dilakukan dengan cara relatif mudah. Secara prinsip DCT melakukan hal yang sama dengan metode sejenis seperti *Fourier Transform* yakni menggunakan fungsi basis sinusoidal. Perbedaannya ada pada penerapan fungsi basis DCT dimana hanya menggunakan fungsi kosinus tanpa melibatkan fungsi sinus. Penentuan fitur didapatkan dari nilai kosinus tertinggi hasil DCT pada sebuah blok piksel tetangga. Besaran blok piksel tetangga dimulai dari  $3 \times 3$  namun untuk lebih detail penjelasan dilakukan pada eksperimen di bab V. Untuk persamaan DCT 2 dimensi pada sebuah citra berukuran  $N \times M$  dapat ditulis:

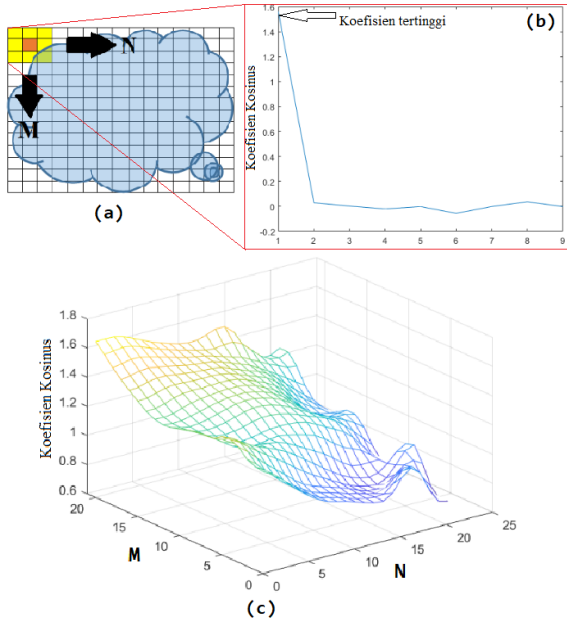
$$C(u, v) = \frac{2}{\sqrt{MN}} \alpha(u) \alpha(v) K \quad (9)$$

dimana nilai  $K$  didapat dari:

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

untuk  $u \in \{1, 2, \dots, N-1\}$  dan  $v \in \{1, 2, \dots, M-1\}$ . Fitur yang akan dipakai didapatkan dari nilai terbesar

hasil DCT untuk mewakili fitur dari blok tetangga piksel.



Gambar. 4 Ilustrasi proses pembentukan fitur DCT. (a) Contoh citra yang akan diproses. (b) Hasil ekstraksi fitur DCT dari 1 blok citra. (c) Hasil seluruh fitur dari citra dengan mengambil koefisien kosinus tertinggi.

Diberikan data citra Gambar. 4(a), maka fitur DCT didapat dari hasil satu blok piksel tetangga sejumlah  $(M - 1)(N - 1)$ . Hasil fitur DCT diambil nilai kosinus tertinggi, dalam hal ini DCT menghasilkan kosinus tertinggi pada index yang pertama seperti terlihat pada Gambar. 4(b). Seluruh fitur pada data citra akan tampak seperti pada Gambar. 4(c).

**Clustering:** Hasil dari fitur DCT selanjutnya dilakukan proses kluster menggunakan *k-means*. Hasil akhir model *background* ditentukan berdasarkan hasil kluster yang bukan merupakan subset *foreground*. Sehingga, apabila terdapat kluster  $C_1, \dots, C_n$  dimana  $n$  adalah jumlah kluster pada *k-means*, maka model *background*  $B_s$  dapat ditulis:

$$B_s = C_n \rightarrow n \notin F_s \quad (10)$$

dimana  $F_s$  adalah model *foreground*.

#### 4.2. Model Foreground

Model *foreground* didapatkan dari hasil proses diferensial pada citra. Hasil yang didapatkan adalah perubahan intensitas tiap piksel pada citra. Apabila dinotasikan  $\nabla I^c$  sebagai gradien dari citra, maka diferensial pertama dari citra dapat dituliskan:

$$\nabla I^c = \begin{bmatrix} \frac{\partial I^c}{\partial x} \\ \frac{\partial I^c}{\partial y} \end{bmatrix} = \begin{bmatrix} F_h^c \\ F_v^c \end{bmatrix} \quad (11)$$

dimana  $\nabla F_h^c$  dan  $\nabla F_v^c$  adalah gradien vektor dari horizontal dan vertikal. Atau dapat dijabarkan menjadi:

$$\nabla F_h^c = \begin{bmatrix} \frac{\partial R}{\partial x} & \frac{\partial G}{\partial x} & \frac{\partial B}{\partial x} \end{bmatrix}$$

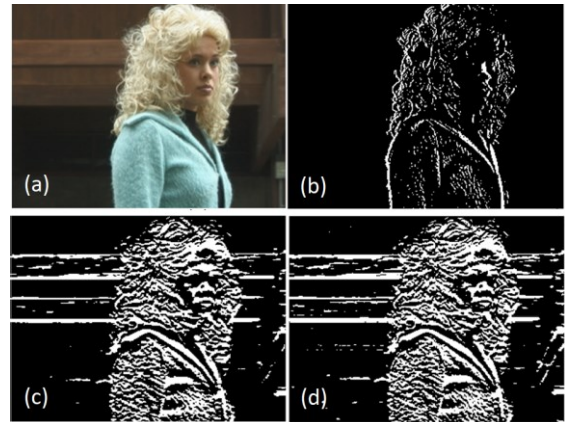
$$\nabla F_v^c = \begin{bmatrix} \frac{\partial R}{\partial y} & \frac{\partial G}{\partial y} & \frac{\partial B}{\partial y} \end{bmatrix}$$

Selanjutnya, mean  $\mu_i^c$  dicari dari setiap gradien pada persamaan (11) sebagai filter noise (*salt dan paper*). Setelah mendapatkan nilai mean serta proses *thresholding*, model *foreground*  $F_s$  dan dapat ditulis menjadi:

$$F_s = \begin{cases} 1, & \text{jika } (\nabla F_h^c - \mu_x^c) + (\nabla F_v^c - \mu_y^c) > \tau \\ 0, & \text{lainnya} \end{cases} \quad (12)$$

dimana  $\tau$  adalah besaran *threshold*

Persamaan (12) berbeda dengan metode fitur *edge detection* pada umumnya. Proses penentuan fitur *edge* dilakukan berdasarkan tiap layer warna (RGB) dengan melakukan proses *background subtraction* terhadap nilai mean. Pada Gambar. 5 menampilkan perbedaan algoritma *edge* yang hasil dari persamaan (12) dengan algoritma *sobel* dan *prewitt*.



Gambar. 5 Perbandingan hasil proses *edge detection* dari beberapa metode. (a) Citra asli. (b) Hasil dari penelitian yang dilakukan. (c) Hasil menggunakan metode *prewitt*. (d) Hasil menggunakan metode *sobel*.

Setelah mengakuisisi kedua model baik *background* dan *foreground*, selanjutnya persamaan (11) dan persamaan (12) disubstitusikan ke dalam variabel  $b_s$  pada persamaan (8) menjadi:

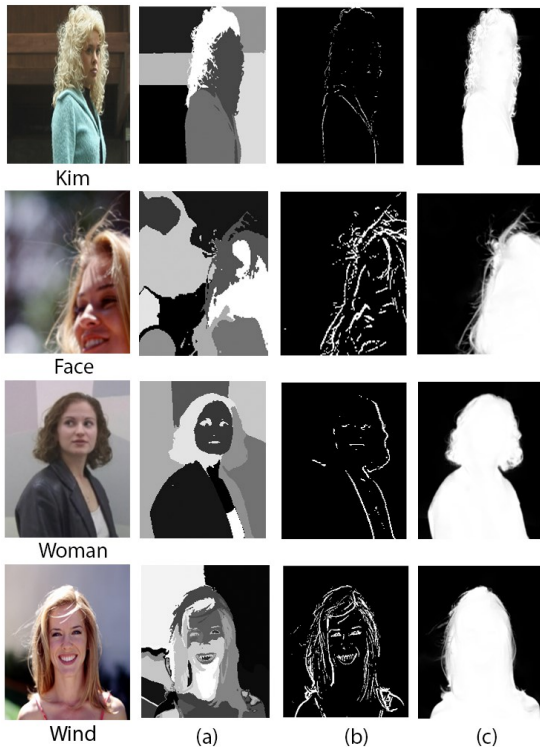
$$(L + \lambda D_s)\alpha = \lambda(F_s + B_s) \quad (13)$$

Oleh karena akuisisi *foreground*  $F_s$  dan *background*  $B_s$  didapat secara otomatis, maka perubahan yang sangat mendasar dari penelitian yang dilakukan adalah merubah metode yang awalnya *supervised* menjadi *unsupervised*.



## 5. HASIL EKSPERIMEN

Pada bab ini akan ditampilkan beberapa hasil dari eksperimen yang telah dilakukan. Seluruh data eksperimen yang digunakan berasal dari data citra yang telah dipakai pada peneliti (Levin, Lischinski and Weiss, 2008). Seluruh eksperimen diproses dan dijalankan pada Intel<sup>(R)</sup>Core<sup>(TM)</sup> i5-8250U @1,6GHz dengan RAM 8GB dan VGA NVIDIA GeForce 940MX (8GB).



Gambar. 6 Hasil penelitian yang telah dilakukan. (a) Akuisisi background. (b) Hasil akuisisi foreground. (c) Hasil alpha matte.

Hasil dari akuisisi foreground dan background secara keseluruhan ditampilkan seperti pada Gambar. 6 data citra Kim, Face, Woman dan Wind dapat menghasilkan alpha matte menggunakan constraint yang telah dibangun. Beberapa nilai parameter dalam eksperimen ditentukan untuk mendapatkan alpha matte terbaik yakni jumlah klaster pada *k-means* pada data citra Kim sebesar 8, data citra Face membutuhkan jumlah klaster 10, untuk Woman dan Wind berturut – turut membutuhkan jumlah klaster 10 dan 40.

Dalam skenario selanjutnya, dilakukan perbandingan kecepatan proses dari metode yang telah dibuat dengan metode *closed-form matting* (Levin, Lischinski and Weiss, 2008) dan *spectral matting* (Levin, Rav-acha and Lischinski, 2008) seperti yang diperlihatkan pada Table 1.

Selanjutnya, eksperimen yang dilakukan adalah membandingkan akurasi hasil alpha matte dari metode yang digunakan dengan alpha matte pada kedua metode sebelumnya. Ground Truth yang dipakai merupakan hasil alpha matte dari metode supervised.

Tabel 1. Perbandingan kecepatan proses

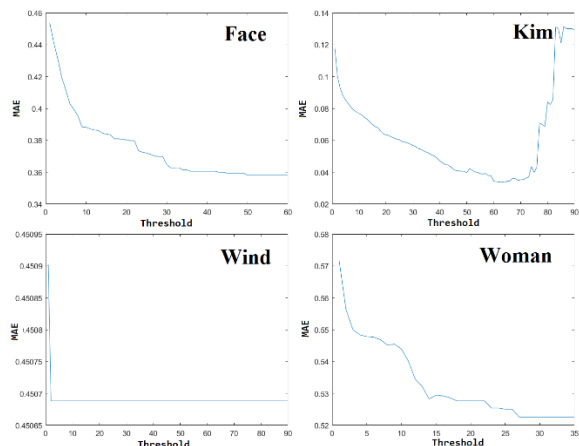
Data Citra (Dimensi)	Waktu yang diperlukan (dalam detik)		
	Spectral	Closed Form	Metode yang diusulkan
Kim (238 x 318)	70,5366	6,8906	7,3321
Face (165 x 137)	25,9910	5,2498	5,3469
Woman (212 x 165)	35,9130	8,0105	8,1338
Wind (254 x 254)	113,1572	7,4317	7,9143

Pengukuran akurasi menggunakan MAE seperti halnya yang dilakukan pada (Zhang *et al.*, 2015) dan (Cho *et al.*, 2016). Persamaan MAE dapat dituliskan dengan:

$$MAE = \frac{1}{MN} \sum_{x=1}^N \sum_{y=1}^M |\alpha(x,y) - \bar{\alpha}(x,y)| \quad (13)$$

dimana  $\bar{\alpha}$  adalah *alpha matte* diperoleh dari metode *Closed Form Matting* (Levin, Lischinski and Weiss, 2008).

Pada Gambar. 7 ditampilkan perbandingan secara quantitative menggunakan MAE. Hasil terbaik menghasilkan nilai MAE minimum.



Gambar. 7 Perbandingan Quantitative: Nilai MAE perbandingan alpha matte antara metode yang diusulkan dengan metode Closed Form.

## 6. ANALISA EKSPERIMEN

### 6.1. Analisa Akuisisi Model

Dalam akuisisi background, beberapa hal yang sangat berpengaruh terhadap hasil eksperimen yaitu penentuan jumlah blok tetangga dan jumlah klaster. Dimensi dari citra yang akan diproses dipengaruhi dari penentuan jumlah blok tetangga. Semakin besar blok tetangga maka semakin besar pula dimensi citra yang akan diproses berkurang. Apabila jumlah blok tetangga diberikan nilai 1, maka dimensi citra akan berkurang menjadi N-2 dan M-2 karena posisi pencarian fitur DCT dimulai dari x+1 sampai N-1 dan y+1 sampai M-1.



Gambar. 8 Hasil kluster menggunakan *k-means* dari citra Wind.  
(a) Hasil *k-means* dengan kluster 10. (b) Hasil *k-means* dengan kluster 20. (c) Hasil *k-means* dengan kluster 30. (d) Hasil *k-means* dengan kluster 10.

Selanjutnya, pengaruh dari jumlah kluster pada tiap data yang diproses berpengaruh terhadap hasil *matting*. Proses kluster dalam penelitian ini berfungsi untuk membantu mengakuisisi *foreground* dan *background*. Karena selain hasil dari kluster *foreground* masih ditentukan oleh subset dari hasil deteksi tepi. Sebagai contoh pada Gambar. 8(a), menunjukkan proses kluster menghasilkan daerah *foreground* menjadi satu dengan daerah *background* pada tempat yang sama. Hal yang sama juga terlihat pada Gambar. 8(b) dan (c) yakni masih ada kluster daerah *foreground* yang masih menjadi satu dengan daerah *background*.

Demikian juga dengan nilai *threshold* pada proses deteksi tepi sangat berpengaruh terhadap hasil *matting*. Penentuan nilai parameter jumlah kluster dan *threshold* menjadi krusial untuk menghasilkan *alpha matte* yang diinginkan. Penelitian ini masih sangat memungkinkan untuk dilanjutkan salah satunya dalam penentuan parameter secara *robust*. Sehingga proses pemilihan nilai parameter baik kluster maupun *threshold* dapat dilakukan juga secara lebih efisien.

## 6.2. Analisa Akurasi dan Kecepatan

Dalam hal kecepatan, nilai perbandingan antara metode penelitian yang dilakukan dengan metode *supervised* tidak jauh berbeda. Perbandingan selisih waktunya tidak lebih dari satu detik. Perbandingan dengan metode *supervised* sangat penting dalam penelitian karena menunjukkan bahwa dengan selisih waktu kurang dari satu detik, metode yang dilakukan dapat mengurangi usaha lebih dalam memberikan label secara manual.

Namun tidak demikian apabila dibandingkan dengan metode *unsupervised* dimana menghasilkan perbandingan kecepatan yang jauh berbeda. Hal ini

disebabkan metode *unsupervised* membutuhkan memory yang cukup besar sehingga membutuhkan waktu yang lebih lama dalam prosesnya. Apabila dibandingkan dengan metode *supervised*, maka penggunaan memory relatif lebih kecil sehingga prosesnya dapat dilakukan relatif lebih cepat.

Pemilihan *ground truth* serta data citra sengaja dipilih dari metode *supervised* sebelumnya. Tujuannya adalah untuk mengevaluasi seberapa besar perubahan hasil *alpha matte* metode *supervised* sebelum dimodifikasi dan sesudah diubah menjadi otomatis. Hasilnya terbukti dengan menunjukkan hasil nilai MAE maksimum 0,336 pada data citra Face dimana secara visual masih dapat ditoleransi.

## 7. KESIMPULAN

Pada penelitian ini memperkenalkan metodologi baru mengenai *image matting* dengan merubah metode *supervised* menjadi *unsupervised*. Penentuan *scribbles* sebagai *constraint* pada metode *supervised matting* yang ditentukan dari akuisisi *foreground* dan *background* berbasis fitur DCT secara otomatis. Penentuan jumlah parameter blok tetangga pada DCT, besaran *threshold* dan jumlah kluster pada *k-means* yang ideal ditentukan oleh varian data pada citra. Hasil perbandingan *alpha matte* dari penelitian yang telah dilakukan dengan metode *supervised* hanya berbeda maksimum sebesar 0,4 detik dari sisi kecepatan dan 0,336 pada skor MAE dan hal ini masih dapat ditoleransi secara visual.

## DAFTAR PUSTAKA

- AKSOY, Y. *et al.* (2018) 'Semantic soft segmentation', *ACM Transactions on Graphics*, 37(4), pp. 1–13. doi: 10.1145/3197517.3201275.
- CHO, D. *et al.* (2016) 'Automatic Trimap Generation and Consistent Matting for Light-Field Images', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PP(99), pp. 1–14. doi: 10.1109/TPAMI.2016.2606397.
- CHO, H.-W. *et al.* (2017) 'Image Matting for Automatic Target Recognition', *IEEE Transactions on Aerospace and Electronic Systems*, 13(9), pp. 1–1. doi: 10.1109/TAES.2017.2690529.
- EISEMANN, M., WOLF, J. AND MAGNOR, M. (2009) 'Spectral video matting', *Proc. VMV*, 11.
- GRADY, L. *et al.* (2005) 'Random walks for interactive alpha-matting', *Proceedings of VIIP*, 2005, pp. 423–429. Available at: [http://cs.tum.edu/fileadmin/user\\_upload/Lehrstuehle/Lehrstuhl\\_XV/Research/Publications/2005/VIIP05.pdf](http://cs.tum.edu/fileadmin/user_upload/Lehrstuehle/Lehrstuhl_XV/Research/Publications/2005/VIIP05.pdf).

- KOESHARDIANTO, M. (2016) 'Video Object Extraction Using Feature Matching Based on Nonlocal Matting', *Conference proceeding: International Seminar on Intelligent Technology and Its Application*, (July), p. 201.
- LEVIN, A., LISCHINSKI, D. AND WEISS, Y. (2008) 'A closed-form solution to natural image matting', *Pami*, 30(2), pp. 228–42. doi: 10.1109/TPAMI.2007.1177.
- LEVIN, A., RAV-ACHA, A. AND LISCHINSKI, D. (2008) 'Spectral matting', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(10), pp. 1699–1712. doi: 10.1109/TPAMI.2008.168.
- LI, W. *et al.* (2016) 'Roto++: Accelerating Professional Rotoscoping using Shape Manifolds', *ACM Transactions on Graphics*, 35(4), pp. 1–14. doi: 10.1145/2897824.2925973.
- LONG, Z., LIU, Y. AND HAN, S. (2017) 'Video segmentation based on strong target constrained video saliency', *2017 2nd International Conference on Image, Vision and Computing, ICIVC 2017*, pp. 356–360. doi: 10.1109/ICIVC.2017.7984577.
- LU, Y. *et al.* (2016) 'Coherent Parametric Contours for Interactive Video Object Segmentation', *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 642–650. doi: 10.1109/CVPR.2016.76.
- MASHTALIR, S. AND MASHTALIR, V. (2016) 'Sequential temporal video segmentation via spatial image partitions', *Proceedings of the 2016 IEEE 1st International Conference on Data Stream Mining and Processing, DSMP 2016*, (August), pp. 239–242. doi: 10.1109/DSMP.2016.7583549.
- MIKSIK, O. *et al.* (2017) 'ROAM: A rich object appearance model with application to rotoscoping', *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017-Janua(2014), pp. 7426–7434. doi: 10.1109/CVPR.2017.785.
- PORTER, T. AND DUFF, T. (1984) 'Compositing digital images', *ACM SIGGRAPH Computer Graphics*, 18(3), pp. 253–259. doi: 10.1145/964965.808606.
- ROTHER, C., KOLMOGOROV, V. AND BLAKE, A. (2004) "GrabCut": interactive foreground extraction using iterated graph cuts', *ACM Transactions on Graphics*, 23(3), p. 309. doi: 10.1145/1015706.1015720.
- SHEN, X. *et al.* (2016) 'Deep automatic portrait matting', *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9905 LNCS, pp. 92–107. doi: 10.1007/978-3-319-46448-0\_6.
- SMITH, A. R. AND BLINN, J. F. (1996) 'Blue screen matting', *SIGGRAPH '96 Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pp. 259–268. doi: <http://dx.doi.org/10.1145/237170.237263>.
- SUN, J. *et al.* (2004) 'Poisson matting', *ACM Transactions on Graphics*, 23(3), p. 315. doi: 10.1145/1015706.1015721.
- TAN, W. *et al.* (2013) 'Automatic Matting of Identification Photos', in *2013 International Conference on Computer-Aided Design and Computer Graphics*. IEEE, pp. 387–388. doi: 10.1109/CADGraphics.2013.60.
- VICAS, G. *et al.* (2013) 'Automatic trimap generation for digital image matting', *2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA 2013*. doi: 10.1109/APSIPA.2013.6694178.
- WANG, J. AND COHEN, M. F. (2007) 'Image and Video Matting: A Survey', *Foundations and Trends® in Computer Graphics and Vision*, 3(xx), pp. 97–175. doi: 10.1561/06000000019.
- ZENG, Z. *et al.* (2017) 'Pixel Modeling Using Histograms Based on Fuzzy Partitions for Dynamic Background Subtraction', *IEEE Transactions on Fuzzy Systems*, 25(3), pp. 584–593. doi: 10.1109/TFUZZ.2016.2566811.
- ZHANG, Y. *et al.* (2015) 'An Automatic Method for Image Matting Based on Saliency Detection \*', 10, pp. 3571–3578. doi: 10.12733/jcis14117.
- ZHU, Q. *et al.* (2015) 'Targeting accurate object extraction from an image: A comprehensive study of natural image matting', *IEEE Transactions on Neural Networks and Learning Systems*, 26(2), pp. 185–207. doi: 10.1109/TNNLS.2014.2369426.