

Automatic Image Annotation Menggunakan Metode Block Truncation dan K-Nearest Neighbor

Duman Care Khrisne¹, Darma Putra²

¹STIKI, Bali

²Teknologi Informasi, Universitas Udayana, Bali

e-mail: duman.lx14@gmail.com¹, ikgdarmaputra@gmail.com²

Abstrak

Sistem temu kembali citra digital berbasis text sangat bergantung pada label dari gambar digital. Dalam penelitian ini, diterapkan gabungan beberapa metode untuk pelabelan sebuah gambar secara otomatis, istilah yang sering digunakan adalah automatic image annotation, teknik ini digunakan untuk menghasilkan label pada gambar agar dapat melakukan pencarian dengan semantik yang diambil dari objek dalam gambar. Automatic image annotation dimulai dengan melakukan segmentasi terhadap gambar dan untuk setiap segmen gambar dilakukan ekstraksi fitur warna dan tekstur, fitur ini dinormalisasi dan disimpan kedalam basis data untuk data latih, data latih yang telah terkumpul dilatih menggunakan metode learning vector quantization. Bobot yang didapat dari hasil pelatihan digunakan untuk melakukan klasifikasi terhadap segmen gambar ke kosa kata hasil terjemahannya. Hasil dari penelitian ini adalah kesimpulan bahwa automatic image annotation dapat dicapai dengan gabungan metode yang diusulkan dan dapat memberi performa hasil anotasi yang bagus, dimana akurasi sistem adalah 73,26 % saat menggunakan K-NN dengan k = 5.

Kata kunci: automatic image annotation, pelabelan, fitur warna, k-nearest neighbor

Abstract

Labeling of digital images is an important role in digital image retrieval system. In this research, combined methods are utilized to create a label for an image automatically, known as automatic image annotation where this technique is used to generate a label on the image that will help image searching with a more refined semantics. Nowadays, as known the author, there has been no scientific work that combining between block truncation algorithm and k-nearest neighbor. Automatic image annotation begins with feature extraction where these features will be labeled and stored into the database as the training data. Afterward, the k-nearest neighbor method was used to classify the test data using the training set in the database. The results of this study is a system can label the image automatically from one of the theme of the image dataset, the accuracy of the system at most is 73,26% while using K-NN with k = 5.

Keywords: automatic image annotation, labeling, color feature, k-nearest neighbor

1. Pendahuluan

Dalam sepuluh tahun terakhir gambar digital telah mengalami pertumbuhan jumlah yang sangat pesat. Internet foto *sharing* sangat digemari, pada april 2007 *Flickr* yang merupakan salah satu media *online* foto *sharing* telah memiliki 5 juta anggota ter-registrasi dan lebih dari 250 juta gambar [1], walaupun penggunaanya sudah mulai melakukan pelabelan gambar namun sebagian besar gambar digital yang terdapat di Internet masih belum terdokumentasikan. Untuk menggali informasi dari gambar digital yang berjumlah banyak, perlu dibuat sebuah teknik untuk mendokumentasikan dan melakukan pencarian kembali terhadap gambar.

Teknik *image retrieval* telah dikembangkan sejak tahun 1970 [2]. Para peneliti dari dua komunitas yang berbeda yaitu, komunitas *database management* dan komunitas *computer vision* menggunakan dua jenis pendekatan yang berbeda untuk melakukan *image retrieval*, *text-*

based dan *visual-based*. *Text-based image retrieval* pada jaman itu mengharuskan gambar di-anotasi secara manual sebelum gambar tersebut dapat diambil atau dicari kembali. Ada dua hal yang menyebabkan pendekatan ini tidak mungkin dilakukan sekarang, pertama banyaknya tenaga dan sumber daya yang digunakan untuk meng-anotasi gambar digital yang banyak saat ini, dan subjektivitas dari orang yang melakukan anotasi. Orang yang berbeda menginterpretasikan gambar dengan cara yang berbeda dan menghasilkan label yang berbeda [3].

Awal tahun 1990 *content-based image retrieval* (CBIR), melakukan pendekatan dengan teknik baru yaitu dengan melakukan *image retrieval* berdasarkan isi gambar secara visual, seperti warna dan tekstur dan tidak menggunakan *keyword* sebagai acuan. Teknik ini mendapat lebih banyak perhatian dibandingkan teknik sebelumnya, namun terjadi masalah, karena sebagai pengganti *keyword*, sebuah gambar harus dijadikan acuan untuk melakukan *image retrieval*. Hal ini menyebabkan suatu kejadian yang disebut dengan *semantic gap*, yaitu kurangnya kemampuan seseorang mendapatkan informasi yang diekstrak dari sebuah data visual yang dimiliki, karena data yang dapat diekstrak dari data visual diinterpretasikan berbeda oleh *user* [3].

Content-based image retrieval dan *text-based image retrieval* memiliki kelemahan dalam proses temu kembali citra digital. Oleh karena itu penelitian *automatic image annotation* hadir sebagai jembatan yang mengatasi kelemahan dari kedua metode tersebut dan pada penelitian ini dilakukan perancangan *automatic image annotation* menggunakan gabungan metode *block truncation algorithm* untuk melakukan ekstraksi fitur *k-nearest neighbor* (K-NN) untuk mengklasifikasi vektor fitur. Tujuannya adalah mengatasi kelemahan *content-based image retrieval* dan *text-based image retrieval*, dengan cara memberikan label yang dibentuk dari informasi dalam gambar digital, secara otomatis pada sebuah gambar digital. Sehingga proses temu kembali citra digital dapat dilakukan dengan pencarian semantik yang lebih baik, karena *keyword* atau label dalam sebuah gambar diekstrak dari ciri yang dimiliki oleh gambar tersebut.

2. Penelitian Sebelumnya

Jiayu Tang pada tahun 2008 telah melakukan penelitian dan membandingkan beberapa pendekatan salah satunya dengan menggunakan metode *cross media relevance model* pada *salient region* yang dibandingkan dengan *cross media relevance model* pada *region based* yang ditulis oleh Jonathon S. Hare and Paul H[3]. Lewis pada penelitian berjudul *image retrieval using salient regions with vector spaces and latent semantics* pada tahun 2005. Data set yang digunakan secara acak dibagi menjadi 3 bagian, 45% digunakan untuk *training set*, 5% digunakan untuk *evaluation set* sedangkan 50% sisanya digunakan untuk *test set*. Didapatkan kesimpulan bahwa *cross media relevance model* pada *salient region* mampu memprediksi kata yang lebih tepat untuk sebuah gambar, tingkat akurasinya sampai dengan 80% namun hanya bagus untuk sebuah kata dalam sebuah gambar, jika jumlah kata yang harus diprediksi dalam sebuah gambar bertambah maka tingkat akurasinya menurun.

Trong-Tôn Pham dalam penelitiannya melakukan penggabungan antara *region-based* and *saliency-based models* untuk melakukan *automatic image annotation* pada tahun 2006 [4]. Hasil uji didasarkan pada kemampuan metode mendapatkan kata untuk diprediksi, sebuah kata dianggap dapat diprediksi jika rata-rata pemanggilan kembali lebih besar dari 0, sedangkan jika tidak maka dianggap tidak dapat diprediksi. Berdasarkan ketentuan tersebut dari 260 kata dalam *test set* didapatkan hasil 87 kata dapat diprediksi dengan *region-based model*, menggunakan *direct-fusion model* mendapatkan 81 kata yang dapat diprediksi, 75 kata terprediksi dengan menggunakan *latent semantic analysis*, sedangkan menggunakan *saliency-based* hanya mampu menghasilkan 36 kata terprediksi.

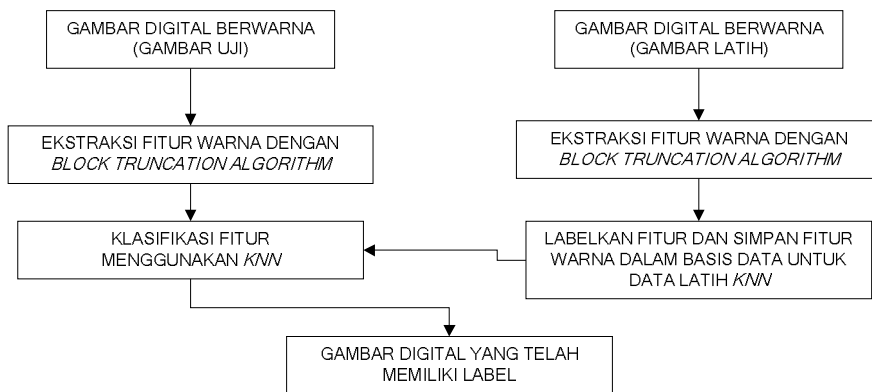
Dr. Sanjay Silakari, Dr. Mahesh Motwani dan Manish Maheshwari dalam penelitiannya melakukan penelitian dengan metode *block truncation algorithm* dalam melakukan temu kembali gambar [5], hasil yang didapatkan menunjukkan bahwa teknik *block truncation algorithm* memberikan presisi yang lebih baik daripada menggunakan momen warna saja. Penelitian ini juga menyimpulkan bahwa isi dari sebuah gambar digital dapat direpresentasikan

dalam fitur-fitur seperti warna, tekstur dan bentuk jika algoritma klusterisasi *k-means* diaplikasikan pada fitur-fitur ini.

Ameesh Makadia, Vladimir Pavlovic dan Sanjiv Kumar telah melakukan penelitian dengan menggunakan 2 pendekatan yaitu *joint equal contribution (JEC)* dan *L1-penalized logistic regression (Lasso)* pada tahun 2008 [6]. Kesimpulan yang didapat dari penelitian ini adalah presisi yang mengejutkan dari teknik *baseline* yang dianggap sederhana namun mampu menyamai metode-metode yang lebih kompleks. Penelitian ini berhasil membuat sebuah pendekatan simple yang menggunakan fitur warna dan tekstur yang didapatkan dari momen *histogram* serta gelombang singkat *Haar* dan *Gabor* serta teknik *K-NN* digunakan dalam proses klasifikasinya.

3. Metode Yang Diusulkan

Penelitian ini memiliki dua bagian utama yaitu proses ekstraksi fitur warna dari data latih menggunakan *block truncation algorithm (BTC)* dan proses melakukan pelabelan dengan menggunakan metode *K-NN*. Skema dari sistem yang diusulkan dapat dilihat pada Gambar 1.



Gambar 1. Skema sistem yang diusulkan

Sistem yang akan dikembangkan menggunakan 2 metode yang digabungkan yang akan dibahas pada subbab 3.1 dan 3.2.

3.1 Block Truncation Algorithm

Block truncation algorithm (BTC) adalah algoritma mendapatkan fitur warna dari sebuah citra berwarna, citra berwarna dibagi berdasarkan komponen penyusun warna R, G dan B, rata-rata dari setiap komponen warna dijadikan patokan untuk memisah komponen warna menjadi dua H dan L dimana H untuk *pixel-pixel* dalam citra yang memiliki nilai lebih tinggi dari rata-rata nilai *pixel* dalam suatu komponen warna dan L untuk *pixel-pixel* dalam citra yang memiliki nilai lebih rendah dari rata-rata nilai *pixel* dalam suatu komponen warna. Jadi warna dari sebuah gambar membentuk 6 kelompok RH, RL, GH, GL, BH dan BL . Momen-momen dari kelompok inilah yang menjadi fitur warna dari *BTC* [5]. Dalam penelitian ini akan digunakan dua buah momen yaitu :

- a. **Mean.** Mean dapat diartikan sebagai rata-rata nilai warna yang terdapat pada gambar digital, momen *mean* dapat dicari dengan Persamaan 1.

$$E_k = \frac{1}{PQ} \sum_{i=1}^P \sum_{j=1}^Q P^{k_{ij}} \quad (1)$$

E = Momen
 k = Komponen warna

- p_{ij}^k = Nilai *pixel* pada (i,j) dalam sebuah komponen warna k
- P = Tinggi gambar
- Q = Lebar gambar

b. Standar Deviasi. Standar deviasi adalah nilai akar kuadrat dari distribusi varian, momen standar deviasi dapat dicari dengan Persamaan 2.

$$SD_k = \sqrt{\frac{1}{PQ} \sum_{i=1}^P \sum_{j=1}^Q (p_{ij}^k - E_k)^2} \dots\dots\dots (2)$$

- SD = Standar Deviasi
- k = Komponen warna
- p_{ij}^k = Nilai *pixel* pada (i,j) dalam sebuah komponen warna k
- E_k = Nilai *mean* pada komponen warna k
- P = Tinggi gambar
- Q = Lebar gambar

3.2 K-Nearest Neighbor

K-nearest neighbor(K-NN) adalah jenis metode klasifikasi non parametrik, yang berarti metode ini tidak memperhatikan distribusi dari data yang ingin kita kelompokkan. Teknik ini sangat sederhana dan mudah untuk diimplementasikan, algoritma dari *K-NN* adalah sebagai berikut [7].




1. Mulai.
2. Input berupa : Data Latih, label untuk data latih, k , data uji.
3. Hitung jarak data uji ke setiap data training.
4. Pilih k data latih yang jaraknya paling dekat dengan data uji.
5. Periksa label dari k data latih yang jaraknya paling dekat dengan data uji.
6. Tentukan label yang frekuensinya paling banyak.
7. Labelkan data uji dengan label yang frekuensinya paling banyak.
8. Stop.




Untuk menghitung jarak antara data uji dan data latih dapat digunakan jarak *Euclidean*.

$$d(x, y) = \|x - y\|^2 = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \dots\dots\dots (3)$$

Dua buah gambar yang memiliki tema yang sama atau gambar yang dilabelkan dengan label gambar yang sama memiliki fitur warna dengan jarak *Euclidean* yang bernilai relatif kecil jika dibandingkan dengan gambar yang tidak memiliki tema atau label gambar yang sama. Tabel 1 membuktikan bahwa gambar yang dilabelkan dengan label yang sama akan memiliki jarak yang lebih kecil jika dibandingkan dengan gambar yang dilabelkan dengan label yang berbeda.

Tabel1. Perbandingan jarak fitur warna terhadap gambar acuan


Gambar	 (acuan)		 (uji 1)		 (uji 2)	
	Fitur Warna	M – RL	21.81276448 56771	M – RL	13.945058186 849	M – RL
	M – RH	89.47943115 23438	M – RH	84.897745768 2292	M – RH	76.651311 3720064

	M – GL	39.79763793 94531	M – GL	35.905537923 1771	M – GL	32.526725 9446151
	M – GH	89.00500488 28125	M – GH	72.360788981 1198	M – GH	81.966498 4637923
	M – BL	28.63138834 63542	M – BL	21.731028238 9323	M – BL	33.998931 7761003
	M – BH	66.52960205 07812	M – BH	50.730082194 0104	M – BH	87.748926 6893198
	SD – RL	28.32713989 7078	SD – RL	18.789773444 1592	SD – RL	36.165092 8673259
	SD – RH	45.63636282 38834	SD – RH	48.647543187 6674	SD – RH	69.526094 6237467
	SD – GL	39.92827450 81015	SD – GL	29.867275312 4444	SD – GL	33.988356 2170152
	SD – GH	54.36267611 57691	SD – GH	63.520961753 7918	SD – GH	68.226359 0709167
	SD – BL	27.88844248 23128	SD – BL	20.672400869 1103	SD – BL	35.044393 1234214
	SD – BH	31.72561762 6994	SD – BH	20.060846984 7561	SD – BH	76.366856 155085
Label	Kuda		Kuda		Gunung dan Glasier	
Jarak ke (acuan)	0		33.84635		62.40925	
Gambar						
	(uji3)		(uji 4)		(uji 5)	
	M – RL	14.15962847 28934	M – RL	4.6231689453 125	M – RL	15.451538 0859375
M – RH	95.09897555 36792	M – RH	68.221750895 1823	M – RH	195.83075 9684245	
M – GL	30.86863281 68713	M – GL	4.3232014973 9583	M – GL	13.747090 6575521	
M – GH	91.95103614 55589	M – GH	13.289774576 8229	M – GH	192.23825 0732422	
M – BL	21.94975431 60015	M – BL	9.3480224609 375	M – BL	42.503356 9335938	
M – BH	63.58261188 03219	M – BH	16.255828857 4219	M – BH	147.89525 3499349	
Fitur Warna	SD – RL	25.15114439 7055	SD – RL	6.9671005494 8611	SD – RL	39.388007 1101041
	SD – RH	39.40641195 8317	SD – RH	49.050524389 6089	SD – RH	28.735147 9993196
	SD – GL	37.93446584 33629	SD – GL	4.5879930757 4405	SD – GL	35.781631 7191425
	SD – GH	51.11111812 18775	SD – GH	18.795063036 8458	SD – GH	30.202775 087356
	SD – BL	25.14262697 0656	SD – BL	6.5378896368 77	SD – BL	60.444189 5812441
	SD – BH	17.15535020 29348	SD – BH	19.718967481 4224	SD – BH	25.324316 7847198
	Label	Kuda		Bunga		Dinosaurus
Jarak ke (acuan)	22.69723		119.2468		177.7902	

4. Hasil

Pengujian terhadap sistem dilakukan menggunakan data latih berjumlah 500 buah data latih yang telah diekstrak fitur warnanya dan dilabelkan sesuai dengan tema, gambar digital yang digunakan diambil dari 'http://wang.ist.psu.edu/~jwang/test1.tar', dataset yang sama digunakan oleh [8] untuk melakukan penelitian dalam bidang pelabelan gambar otomatis. File 'test1.tar' berisi 1000 buah gambar digital yang dibagi menjadi 10 buah tema yaitu Orang Afrika dan Desa, Pantai, Gedung atau Bangunan, Bus, Dinosaur, Gajah, Bunga, Kuda, Gunung dan Glasier dan Makanan, dari 1000 buah gambar digital yang tersedia 500 buah gambar dijadikan gambar latih dan sebanyak 90 gambar akan digunakan sebagai data uji.

Analisis Gambar



kelas	kelas1
Orang Afrika dan Desa	12.65972910481...
Orang Afrika dan Desa	13.08953244712...
Gajah	14.30163667610...
Gajah	14.70327802364...
Orang Afrika dan Desa	18.29086598753...
Orang Afrika dan Desa	18.39348504605...
Gajah	19.864940537178
Gajah	20.13674885275...
Gajah	20.13858419054...
Orang Afrika dan Desa	20.58373782677...

Jumlah K : 3 Tentukan Tema Gambar Tema Gambar : Orang Afrika dan Desa

17.3704310318518,101.727545549802,34.6941779672224,84.7669966123765,28.0972644686111,60.9706711156777

(a)

Analisis Gambar



kelas	kelas1
Orang Afrika dan Desa	12.65972910481...
Orang Afrika dan Desa	13.08953244712...
Gajah	14.30163667610...
Gajah	14.70327802364...
Orang Afrika dan Desa	18.29086598753...
Orang Afrika dan Desa	18.39348504605...
Gajah	19.864940537178
Gajah	20.13674885275...
Gajah	20.13858419054...
Orang Afrika dan Desa	20.58373782677...

Jumlah K : 5 Tentukan Tema Gambar Tema Gambar : Orang Afrika dan Desa

17.3704310318518,101.727545549802,34.6941779672224,84.7669966123765,28.0972644686111,60.9706711156777

(b)

Gambar 2. Hasil pelabelan gambar uji dengan (a) k bernilai 3, (b) k bernilai 5

Gambar uji diinputkan kedalam sistem dan sistem akan menghitung jarak terdekat dari fitur data uji ke semua fitur data latih yang terdapat dalam basis data, dengan menggunakan metode K - NN . Dengan menginputkan nilai k dicari k nilai terdekat dari jarak data uji yang ingin diklasifikasikan, kelas yang memiliki frekuensi paling banyak akan menjadi label dari gambar yang diinputkan. Gambar 2 menunjukkan hasil pelabelan gambar dengan k bernilai 3 dan k bernilai 5.

Gambar uji yang berjumlah 90 buah gambar akan diuji pelabelannya dengan variasi nilai k sehingga mendapatkan hasil seperti yang dapat dilihat pada Tabel 2.

Tabel 2. Akurasi pelabelan dengan variasi jumlah k pada k -NN

Nilai k	Label benar	Label Salah	Akurasi
3	61	27	67,78%
5	63	23	73,26%
10	58	32	64,44%

Nilai k sangat berpengaruh terhadap hasil pelabelan yang dilakukan oleh sistem, Tabel 2 memberikan informasi bahwa pelabelan dengan nilai $k = 5$ memberi akurasi pelabelan yang lebih baik jika dibandingkan dengan pelabelan dengan nilai $k = 3$ atau $k = 10$.

5. Simpulan

Block truncation algorithm sebagai salah satu algoritma ekstraksi fitur warna telah mampu melakukan ekstraksi fitur yang dapat memisahkan ciri gambar dari sepuluh tema yang digunakan dalam penelitian. Gabungan metode *block truncation algorithm* dan *k-nearest neighbor* yang diusulkan terbukti mampu melakukan pelabelan gambar secara otomatis dengan akurasi sistem yang dikembangkan mencapai 73,26%. Penggunaan *block truncation algorithm* masih dapat digabungkan dengan metode ekstraksi fitur yang lain, sehingga kedepannya performa kerja dari sistem yang dikembangkan masih dapat ditingkatkan. Sedangkan jika dilihat dari teknik klasifikasinya teknik *k-nearest neighbor* adalah teknik klasifikasi non parametrik, dan masih dapat dikembangkan dan diganti dengan teknik klasifikasi menggunakan jaringan syaraf tiruan.

Daftar Pustaka

- [1] Morgan Ames, Mor Naaman, "Why we tag: motivations for annotation in mobile and online media", Proceedings of the SIGCHI conference on Human factors in computing systems, ACM, pp.971-980, 2007.
- [2] Y. Rui, T. Huang, and S. Chang, "Image retrieval: Current techniques, promising directions and open issues", Journal of Visual Communication and Image Representation, 10(4), pp.39-62, April 1999.
- [3] Jiayu Tang, "Automatic Image Annotation and Object Detection", Thesis, University of Southampton, 2008.
- [4] Trong-Ton Pham, "Automatic Image Annotation: Towards a Fusion of Region-based and Saliency-based Models", Disertasi, Universite Pierre Et Marie Curie Master Iad, 2006.
- [5] Sanjay Silakari, Mahesh Motwani, Manish Maheshwari, "Color Image Clustering using Block Truncation Algorithm", IJCSI International Journal of Computer Science Issues, Vol. 4, No. 2, pp. 31-36, 2009.
- [6] Ameesh Makadia, Vladimir Pavlovic, Sanjiv Kumar, "Baseline for Image Annotation", Google Research New York & Rutgers University Picastaway, 2009.
- [7] Santosa, Budi, "Data Mining: Teknik Pemanfaatan Data unuk Keperluan Bisnis", Graha Ilmu, Yogyakarta, 2007.
- [8] Jia Li, James Z. Wang, "Automatic linguistic indexing of pictures by a statistical modeling approach", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No.9, pp.1075-108, 2003.