

# Relevance Feedback berdasarkan Support Vector Machine pada Content Based Image Retrieval

Virginia Tulenan

Program Studi Teknik Informatika, Universitas Sam Ratulangi  
Jl. Kampus Unsrat Bahu, Manado, Indonesia

virginia.tulenana@unsrat.ac.id

**Abstract**— *Content based image retrieval* adalah bidang penelitian yang sangat penting saat ini dalam bidang multimedia database. Banyak penelitian yang telah dilakukan dalam dekade terakhir untuk merancang teknik *image retrieval* yang efisien dari image database. Meskipun banyak teknik pengindeksan dan *retrieval* telah dikembangkan, namun masih belum terdapat teknik pemisahan ciri (*feature extraction*), *indexing* dan *retrieval* yang bisa diterima secara universal oleh semua orang. Dalam tulisan ini, digunakanlah metode *relevant feedback* berdasarkan *support vector machine (SVM)* dan *muhlabois distance* untuk pengukuran kemiripan pada *image retrieval*.

**Keywords**— *Content based image retrieval, indexing, feature extraction, retrieval, relevant feedback, support vector machine.*

## I. PENDAHULUAN

Content based image retrieval (CBR) telah menjadi bidang penelitian yang sangat aktif sejak 1990an. Beberapa usaha penelitian berfokus pada penemuan fitur gambar terbaik dan cara merepresentasikannya. Salah satunya adalah dengan menjumlahkan jarak dari masing-masing fitur representasi dengan bobot tetap. Tulisan ini memperkenalkan pendekatan berdasarkan *relevance feedback* dimana dalam pendekatan ini user harus tepat menguraikan informasi dalam *weight* (bobot) yang tepat dan memungkinkan user untuk memasukan query awal yang kasar dan secara berkelanjutan memperbaiki kebutuhan informasi dengan umpan balik yang relevan. Pendekatan ini mengurangi upaya user dalam menyusun suatu query dan merekam kebutuhan informasi lebih tepat. Dengan memanfaatkan umpan balik positif dan negatif maka diharapkan persepsi subyektivitas user dapat diambil secara lebih akurat sehingga hasil yang diperoleh pun sesuai dengan keinginan user.

## II. EKSTRAKSI FITUR

### A. Color: Color histogram

Warna adalah fitur yang penting dan paling kelihatan yang diterima/ dirasakan oleh manusia ketika pertama kali melihat sebuah gambar. *Color histogram* adalah salah satu metode yang dipakai untuk merepresentasikan konten dari warna. Alasannya adalah warna dari histogram bersifat independen dari ukuran gambar dan orientasi. Secara statistic, *color histogram* menunjukkan kemungkinan penggabungan dari intensitas 3 channel warna. Dalam metode ini, gambar disegmentasi dan warna diekstrak dari setiap bagian. Data statistic utama yang diekstrak adalah nilai rata-rata, deviasi

standar dan nilai tengah untuk setiap channel warna yaitu RGB (red, green, blue). Jadi total  $3 \times 3 = 9$  fitur per segmen yang diperoleh.

### B. Shape: Wavelet transform

Sebuah gambar dengan ukuran  $M \times N$  dapat dikomposisikan kedalam wavelet koefisien menggunakan Mallat pyramid algorithm [3]. Menurut paper [4] secara matematik dapat diekspresikan kedalam persamaan rekursif sebagai berikut:

$$LL^{(k)}(m, n) = [[LL_{rows}^{(k-1)} * \bar{H}]_{2 \downarrow 1 \text{ columns}} * \bar{H}]_{1 \downarrow 2},$$

$$m = 1, \dots, M/2^k; \quad n = 1, \dots, N/2^k,$$

$$LH^{(k)}(m, n) = [[LL_{rows}^{(k-1)} * \bar{H}]_{2 \downarrow 1 \text{ columns}} * \bar{G}]_{1 \downarrow 2},$$

$$m = M/2^k + 1, \dots, M/2^{k-1}; \quad n = 1, \dots, N/2^k,$$

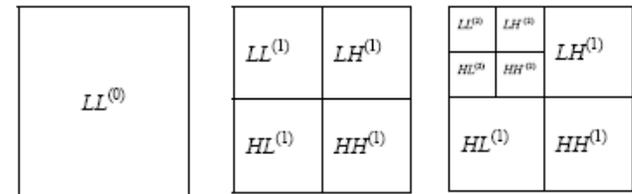
$$HL^{(k)}(m, n) = [[LL_{rows}^{(k-1)} * \bar{G}]_{2 \downarrow 1 \text{ columns}} * \bar{H}]_{1 \downarrow 2},$$

$$m = 1, \dots, M/2^k; \quad n = N/2^k + 1, \dots, N/2^{k-1},$$

$$HH^{(k)}(m, n) = [[LL_{rows}^{(k-1)} * \bar{G}]_{2 \downarrow 1 \text{ columns}} * \bar{G}]_{1 \downarrow 2},$$

$$m = M/2^k + 1, \dots, M/2^{k-1}; \quad n = N/2^k + 1, \dots, N/2^{k-1}.$$

LL, LH, HL dan HH merepresentasikan 4 sub gambar yang didekomposisi. Setelah wavelet didekomposisi, setiap fitur didistribusikan kedalam sebuah subband yang berbeda-beda dengan frekuensi komponen yang spesifik.



(a) (b) (c)

Gbr. 1 Ilustrasi dari wavelet dekomposisi, (a) Dekomposisi level 0, (b) Dekomposisi level 1, (c) Dekomposisi level 2

## III. SUPPORT VECTOR MACHINE (SVM)

SVM merupakan metode *machine learning* yang bekerja atas prinsip *structural risk minimization* (SRM) dengan tujuan menemukan hyperplane terbaik yang memisahkan dua buah kelas pada ruang input. SVM pertama kali diperkenalkan oleh Vapnik pada tahun 1992 sebagai rangkaian harmonis konsep-konsep unggulan dalam bidang *pattern recognition*. Walaupun

usia SVM masih relative muda, evaluasi kemampuannya dalam berbagai aplikasi menempatkannya sebagai *state of the art* dalam *pattern recognition*, dan merupakan salah satu tema yang berkembang dengan pesat.

Data yang tersedia dinotasikan sebagai  $x_i \in R^d$  sedangkan label masing-masing dinotasikan  $y_i \in \{-1, +1\}$  untuk  $i = 1, 2, \dots, /$  dimana  $/$  adalah banyaknya data. Diasumsikan kedua kelas  $-1$  dan  $+1$  dapat terpisah secara sempurna oleh hyperplane berdimensi  $d$  yang didefinisikan  $w \cdot x_i + b = 0$  (1).

Pattern  $x_t$  yang termasuk kelas  $-1$  (sampel negatif  $\alpha_i$ ) dapat dirumuskan sebagai pattern yang memenuhi pertidaksamaan  $w \cdot x_i + b \leq 0$  (2). Sedangkan pattern yang termasuk kelas  $+1$  (sampel positif) dirumuskan  $w \cdot x_i + b \geq 0$  (3).

Margin terbesar dapat ditemukan dengan memaksimalkan nilai jarak antar hyperplane dan titik terdekatnya dan dapat dirumuskan sebagai *quadratic programming* (QP) yaitu mencari titik minimal dengan menggunakan persamaan

$$\text{Min } \tau(w) = \frac{1}{2} \|w\|^2 \quad (4)$$

dengan memperhatikan konstrain

$$y^i (x^i \cdot w + b) - 1 \geq 0, \quad (5)$$

Masalah ini dipecahkan dengan berbagai macam teknik komputasi diantaranya yaitu Lagrange Multiplier

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l \alpha_i (y_i ((x_i \cdot w + b) - 1)) \quad (6)$$

( $i = 1, 2, 3, \dots$ )

adalah Lagrange Multiplier yang bernilai nol atau positif ( $\alpha_i = 0$ ). Nilai optimal dari persamaan (6) dapat dihitung dengan meminimalkan  $L$  terhadap  $w$  dan  $b$  serta memaksimalkan  $L$  terhadap  $\alpha_i$ . Dengan memperhatikan sifat bahwa pada titik optimal gradien  $L = 0$ , persamaan (6) dapat dimodifikasi sebagai maksimalisasi masalah yang hanya mengandung  $\alpha_i$  saja seperti persamaan dibawah ini:

$$\sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j x_i x_j \quad (7)$$

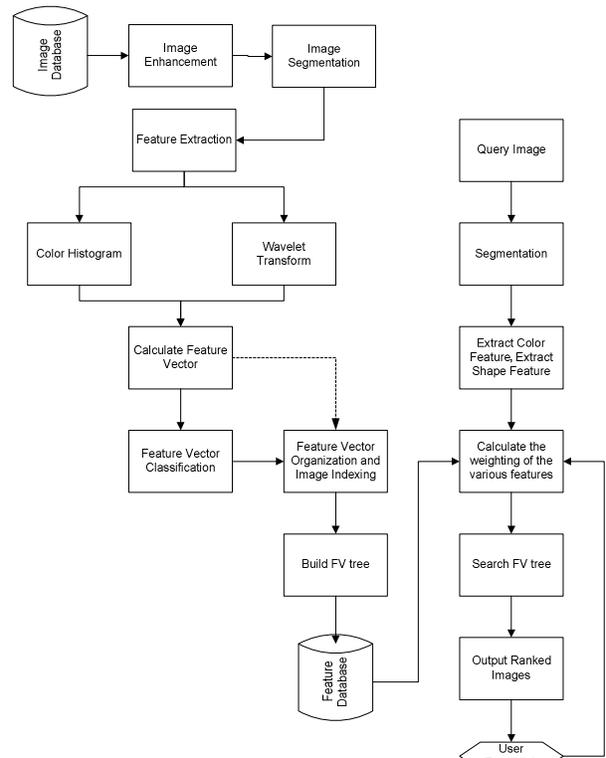
dengan batasan:

$$\alpha_i \geq 0 \quad (i=1, 2, 3, \dots, l) \quad \sum_{i=1}^l \alpha_i y_i = 0 \quad (8)$$

Dari perhitungan ini diperoleh  $\alpha_i$  yang kebanyakan bernilai positif. Data yang berkorelasi dengan  $\alpha_i$  yang positif inilah yang disebut support vector.

#### IV. CBIR FRAMEWORK

Berikut adalah framework *content based image retrieval* yang usulkan.



Gbr. 2 framework CBIR yang diusulkan

#### A. Tahap indexing

Langkah pertama indexing gambar adalah segmentasi. Hasil dari tahap segmentasi biasanya berbentuk komponen gambar seperti sub-gambar, edges, boundary contours, color/intensity measurements, texture measurements dan sebagainya. Teknik segmentasi termasuk variasi active contour segmentation dan active shape modelling. Fitur yang di segmentasi selanjutnya direpresentasikan ke dalam suatu bentuk yang cocok untuk indexing dan komputasi kemiripan. Fitur-fitur tersebut kemudian diklasifikasi dan dilakukanlah image indexing.

#### B. Tahap retrieval

##### 1) Relevance Feedback

Relevance feedback merupakan suatu proses dari secara otomatis menyesuaikan existing query menggunakan information feed-back dari user mengenai relevansi yang diambil dari dokumen sebelumnya. Pada MARS [2], user memberitahu sistem mengenai image mana yang pada hasil retrieved sebelumnya relevant dengan apa yang dia inginkan dan menyediakan pilihan weight untuk tiap relevant image. Query weights untuk tiap low-level feature, seperti, color dan tekstur kemudian secara dinamis diupdate sesuai feedback user. Hal ini membebaskan user dari kesulitan. User tidak lagi membutuhkan untuk menentukan dengan tepat weight untuk tiap low-level feature pada tahap query formulation.

##### 2) Similarity Measurement

Terdapat dua similarity measurement yang dapat digunakan yaitu Mahalanobis distance dan Euclidean distance. Berdasarkan [5] disimpulkan bahwa Mahalanobis distance lebih baik daripada Euclidean distance. Menurut [1] secara keseluruhan similarity  $S_j$  untuk gambar ke- $j$  dalam database diperoleh dengan secara linear menggabungkan individual feature, similarity distance  $f_i$ ,  $S_j(f_i)$  dengan persamaan:

$$S_j = \sum W_i S_j(f_i) \quad j = 1, 2, \dots, N \quad (1)$$

dimana  $N$  merupakan jumlah gambar dalam database.

Muhalunobis distance digunakan sebagai untuk mengukur kemiripan:  $S_j(f_i) = (x_i - q_i)^T C_i^{-1} (x_i - q_i)$   $j = 1, 2, \dots, N$  (2) dimana  $x_i$  merupakan fitur vektor ke- $i$  dari gambar ke  $j$  dalam database,  $q_i$  merupakan fitur vektor dari query, dan  $C_i$  merupakan matrik kovarian dari komponen-komponen fitur ke- $i$  gambar-gambar yang relevan. Elemen dari  $C_i$  ditetapkan sebagai:

$$C_i(m, n) = \frac{\sum_{k=1}^N V(k) [r_i(k, m) - q_i(m)][r_i(k, n) - q_i(n)]}{\sum_{k=1}^N V(k)} \quad (3)$$

dimana  $V(k)$  merupakan preferensi berat dari gambar ke- $k$  yang relevan (*positive feedback*),  $r_i(k, m)$  dan  $r_i(k, n)$  adalah nilai komponen ke- $m$  dan ke- $n$  dari fitur ke- $i$  dari gambar ke- $k$  yang relevan.  $q_i(m)$  dan  $q_i(n)$  adalah nilai komponen ke- $m$  dan ke- $n$  dari ke- $i$  fitur query.  $NR$  adalah banyaknya *positive feedback* dan  $NR > 1$ .  $C_i$  adalah matrik identitas bila  $NR = 1$ .

Proses klasifikasi tidak dapat dilakukan dari *positive feedback* saja. Oleh sebab itu, contoh-contoh negatif masih akan digabungkan dengan contoh-contoh positif setelah *relevance feedback*. Terkadang sulit untuk user menyatakan kepada sistem mengenai subjektifitasnya dengan menentukan pilihan bobot untuk setiap gambar yang relevan. Kesalahan pada query akan terjadi jika pemberian bobot tidak dilakukan dengan benar. Oleh sebab itu digunakanlah klasifikasi preferensi yang lebih lemah seperti suka atau tidak suka sebagai ganti pemberian preferensi bobot yang spesifik. Dalam hal ini, user menyediakan tidak hanya umpan balik yang positif tetapi juga umpan balik yang negatif.

## V. KESIMPULAN

Menyediakan pilihan bobot dari image yang relevan terkadang membebani user dan tidak akurat. Pendekatan Support Vector Machine (SVM) pada Content Based Image

Retrieval (CBIR) membebaskan user dari kesulitan tersebut dan secara otomatis menghasilkan bobot yang diinginkan (*preference weights*) dari gambar-gambar yang relevan. Dalam tulisan ini digunakanlah contoh negative dan positif dari *image retrieval*. Contoh-contoh yang positif adalah gambar yang diinginkan (*preferred image*) sedangkan contoh-contoh yang negatif adalah gambar yang tidak diinginkan (*undesirable image*). Semakin jauh contoh-contoh positif dari hyperplane, makin tidak bisa dibedakan dari contoh-contoh yang negative. Oleh karena itu penentuan bobot harus ditetapkan lebih besar sehingga contoh-contoh negatif akan dihapus dari daftar image yang diterima dan hasil query lebih tepat. SVM merupakan salah satu pendekatan untuk mendapatkan solusi langsung dari proses retrieval yang kompleks. Output dari SVM, bukanlah merupakan nilai probalistik melainkan suatu pengukuran jarak dari suatu hyperplane yang terpisah dalam ruang fitur. Relevance feedback dapat digabungkan untuk menghasilkan query yang lebih tepat sasaran.

## REFERENCES

- [1] Q.Tian, P. Hong, T.S.Huang, "Update Relevant Image Weights for Content-Based Image Retrieval Using Support Vector Machines", IEEE Multimedia and Expo, IEEE International Conference on Volume 2, Issue, Page(s):1199 - 1202 vol.2, 2000
- [2] Y.Rui, T.S. Huang, S. Mehrotra, "Content-Based Image Retrieval With Relevance Feedback in MARS", Proc. IEEE Intl. Conf. on ICP'97, Santa Barbara, CA, 1997
- [3] L.Huang, X.Huang, "Multiresolution Recognition of Offline Handwritten Chinese Characters with Wavelet Transform, " Proc. Int. Conf. Document Analysis and Recognition, pp. 631-634, (2001)
- [4] S.G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," IEEE Trans. On Pattern Analysis and Machine Intelligence, 11(7), pp.674-693, (1989)
- [5] Will Dwinell, <http://matlabdatamining.blogspot.com/2006/11/mahalanobis-distance.htm>
- [6] M,N,R., Soelaiman,R., Helen,A., Karlita,T. Metode Pembelajaran Incremental Support Vector Machine (SVM) Studi Kasus Pengenalan Pola dengan Kelas Banyak. IES (2005)
- [7] Burges,C.J.C. "A tutorial on support vector machines for pattern recognition", Data Mining and Knowledge Discovery 2(2), pp. 121—167 (1998)
- [8] V. Vapnik, The Nature of Statistical Learning Theory. Springer-Verlag, New York, 1995.
- [9] Sanderson and A. Shilton, Image Tracking and Recognition with Applications to Fisheries An Investigation of Support Vector Machines, Final year honours thesis, Department of Electrical and Electronic Engineering, The University of Melbourne, 1999.
- [10] M. Palaniswami, A. Shilton, D. Ralph and B. D. Owen, "Machine Learning using Support Vector Machines," in Proceedings of International Conference on Artificial Intelligence in Science and Technology, AISAT2000, 2000.
- [11] Shilton, M. Palaniswami, D. Ralph and A. C. Tsoi, "Incremental Training of Support Vector Machines," in Proceedings of International Joint Conference on Neural Networks, (IJCNN'), 2001, CD version only.
- [12] Boser, B., Guyon, I., Vapnik, V.: An training algorithm for optimal margin classifiers. In: Fifth Annual Workshop on Computational Learning Theory, Pittsburgh, ACM (1992) pp.144–152