Research Article

Fast Retrieval Algorithm for Earth Mover's Distance Using EMD Lower Bounds and a Skipping Algorithm

Masami Shishibori,[1] Daichi Koizumi,[1] and Kenji Kita[2]

[1] *Department of Information Solution, Institute of Technology and Science, The University of Tokushima,*
 2-1 Minami-Josanjima-Cho, Tokushima-Shi, Tokushima 770-8506, Japan
[2] *Center for Advanced Information Technology, The University of Tokushima, 2-1 Minami-Josanjima-Cho,*
 Tokushima-Shi, Tokushima 770-8506, Japan

Correspondence should be addressed to Masami Shishibori, bori@is.tokushima-u.ac.jp

The earth mover's distance (EMD) is a measure of the distance between two distributions, and it has been widely used in multimedia information retrieval systems, in particular, in content-based image retrieval systems. When the EMD is applied to image problems based on color or texture, the EMD reflects the human perceptual similarities. However, its computations are too expensive to use in large-scale databases. In order to achieve efficient computation of the EMD during query processing, we have developed "fastEMD," a library for high-speed feature-based similarity retrievals in large databases. This paper introduces techniques that are used in the implementation of the fastEMD and performs extensive experiments to demonstrate its efficiency.

1. Introduction

The earth mover's distance (EMD) was introduced in computer vision as an improved distance measure between two distributions, and it has been widely used in multimedia databases. In particular, in the area of content-based image retrieval (CBIR), where color images are retrieved from multimedia databases, it is important to apply improved distance measures. Several CBIR models have been proposed based on the histogram approach, such as the query by image content (QBIC) method proposed by Faloutsos et al. [1], which maps each image to a vector based on encoded attributes such as color distribution, texture, and shape. However, the histogram approaches using Euclidean distance and quadratic forms apply bin-to-bin distance functions; hence, these approaches are sensitive to minor shifts in feature value distribution.

The similarity distance in the EMD is calculated as the amount of changes necessary to transform one image feature into another. Rubner and Tomasi [2] have proposed a CBIR model using the EMD, and its high quality of image-similarity searches has been demonstrated. Shishibori et al. [3]

have proposed a music-retrieval model using the EMD, and a query-by-humming system presents a more robust approach than the conventional methods.

The EMD is defined as a linear programming problem that can be solved using the simplex method. However, computation using the EMD and the simplex method is too complex for its application in interactive multimedia database systems. In order to achieve efficient retrieval processing in large-scale multimedia databases based on the EMD, we propose a fast retrieval algorithm for the EMD.

In order to reduce the computation times of the original distance, the proposed method uses the lower bounding distance described in Cohen and Guibas [4] and Assent et al. [5]. The lower bounding distance of the EMD is a simple approximation in terms of computation. In the case of the k-nearest-neighbor search task, this method calculates the original EMD for the first k images, where the largest distance is set to the threshold. This method calculates the lower bounding distance for $k + 1$ images. This method does not calculate the original EMD if the lower bounding distance exceeds the threshold, because the true EMD cannot be lower than the first k images. In all other cases, this method

calculates the original EMD, and if the true EMD is lower than the first k images, a new image is registered and the threshold is reset.

The remaining sections in this paper are organized as follows: Section 2 introduces the outline of the earth mover's distance. Section 3 explains the lower bounds of the EMD and proposes the algorithm of the calculation skip using the lower bounds. Section 4 evaluates the effectiveness of the presented method by using experimental results. Section 5 describes conclusion and future works.

2. Earth Mover's Distance

2.1. Conventional Dissimilarity Measures. The L_p distance is a generalized measure of the similarity between two histograms $x = \{x_i\}$ and $y = \{y_i\}$. The L_p distance is defined as

$$L_p(x, y) = \left(\sum_{i=0}^{n} |x_i - y_i|^p \right)^{1/p}, \tag{1}$$

where $p > 0$.

The Minkowski distance assumes different names depending on the value of p. If $p = 1$, the distance is called a Manhattan distance. The Euclidean distance can be computed from (1) by selecting $p = 2$. The Euclidean distance is one of the simplest and most popular distance measures.

The L_p distance is a representative of bin-by-bin dissimilarity measures, and it has been used as the the dissimilarity measures of multimedia information retrieval systems. A major disadvantage of this method is that it accounts only for the correspondence between bins with the same index, and it does not use information across bins. This is illustrated in Figure 1, which shows two pairs of one-dimensional gray-scale histograms as described in Rubner and Tomasi [2].

Although the two histograms on the left are the same except for a shift by one bin, the L_1 distance between them is larger than the L_1 distance between the two histograms on the right. As shown in Figure 1(a), the L_p distance does not match the perceptual dissimilarity. This can be fixed using the correspondences between the bins in the two histograms and the ground distance between them, as shown in Figure 1(b). As shown in Figure 1(b), the perceptual dissimilarity is based on the correspondence between the bins in the two histograms.

The quadratic-form distance is a representative of the cross-bin dissimilarity measures. Niblack et al. [6] suggested the use of this distance for color-based retrieval. The quadratic-form distance does not enforce a one-to-one correspondence between the mass elements in the two histograms. The same mass in a given bin of the first histogram is made to correspond to the masses contained in different bins of the other histogram. However, the image-retrieval results using the quadratic-form distance include some false positives, because the quadratic-form distance tends to overestimate the mutual similarity of color distributions, as mentioned in Stricker and Orengo [7].

2.2. Earth Mover's Distance. The EMD is based on the following linear programming problem. Let $P = \{(p_1, w_{p_1}), \ldots,$

$(p_m, w_{p_m})\}$ be the first signature with m clusters, where p_i is the cluster representative and w_{p_i} is the weight of the cluster. Let $Q = \{(q_1, w_{q_1}), \ldots, (q_n, w_{q_n})\}$ be the second signature with n clusters. Let $D = [d_{ij}]$ be the ground-distance matrix, where $d_{ij} = d(p_i, q_j)$ is the ground distance between cluster p_i and q_j.

We assume that the signature P denotes "supply places," the signature Q denotes "demand places," and the ground distance d_{ij} denotes the transportation cost between each supply and demand place. Our aim is to find a flow $F = [f_{ij}]$ such that f_{ij} lies between p_i and q_j, and it minimizes the following overall cost:

$$\text{WORK}(P, Q, F) = \sum_{i=1}^{m} \sum_{j=1}^{n} d_{ij} f_{ij}. \tag{2}$$

Equation (2) is obtained with the following constraints:

$$f_{ij} \geq 0, \quad 1 \leq i \leq m, \ 1 \leq j \leq n, \tag{3}$$

$$\sum_{j=1}^{n} f_{ij} \leq w_{p_i}, \quad 1 \leq i \leq m, \tag{4}$$

$$\sum_{i=1}^{m} f_{ij} \leq w_{q_j}, \quad 1 \leq j \leq n, \tag{5}$$

$$\sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij} = \min\left(\sum_{i=1}^{m} w_{p_i}, \sum_{j=1}^{n} w_{q_j} \right). \tag{6}$$

Constraint (3) allows the movement of supplies from P to Q and not vice versa. Constraint (4) limits the amount of supplies that can be sent by the clusters in P to their weight. Constraint (5) limits the clusters in Q to so that the supplies do not exceed the weights Constraint (6) forces the maximum possible amount of supplies possible to be moved. This amount is called the *total flow*. Once the transportation problem is solved, and we have found the optimal flow F, the EMD is defined as the resulting work normalized by the total flow

$$\text{EMD}(P, Q) = \frac{\text{WORK}(P, Q, F)}{\sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij}}. \tag{7}$$

An example of a two-dimensional EMD is shown. The signature A as supplies is shown in Table 1, and the signature B as demands is shown in Table 2. We assume that the ground distance is the Euclidean distance. The transportation costs between signatures A and B are calculated as shown in Table 3. By solving the transportation problem under the above condition, we can obtain the optimized transportation flow and weight, as shown in Figure 2.

In Figure 2, the white circles indicate supply places, and the black circles indicate demand places. The numbers inside the circle denote the weight of the supplies or demands. The arrows show the transportation flow, and the numbers beside the arrow denote the transportation weight. The total flow cost of Figure 2 is obtained as follows:

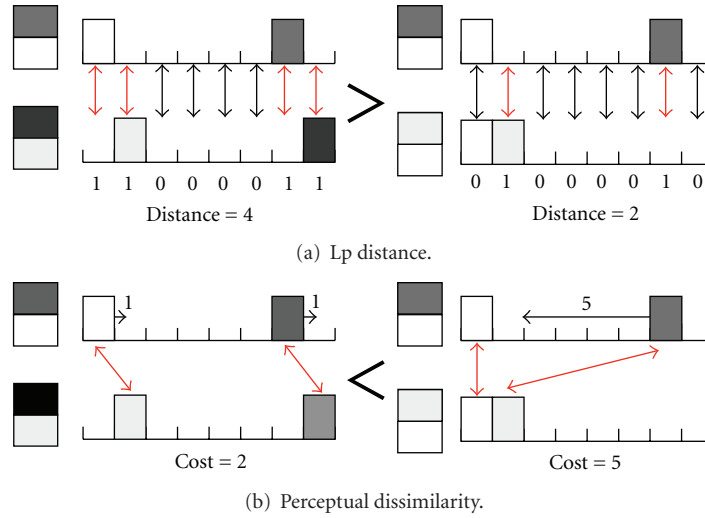$$2 \cdot (0.3 \cdot 1.4142 + 0.2 \cdot 2.6926 + 0.1 \cdot 4.1231) \approx 2.7502. \tag{8}$$

(a) Lp distance.

(b) Perceptual dissimilarity.

FIGURE 1: An illustration of perceptual dissimilarity.

TABLE 1: An example of the signature A.

	Weight	x	y
A_0	0.3	5	5
A_1	0.3	0	5
A_2	0.3	0	0
A_3	0.3	5	0

TABLE 2: An example of the signature B.

	Weight	x	y
B_0	0.4	1	1
B_1	0.4	4	1
B_2	0.4	2.5	4

TABLE 3: Cost matrix of signature A and B.

	B_0	B_1	B_2
A_0	5.6569	4.1231	2.6926
A_1	4.1231	5.6569	2.6926
A_2	1.4142	4.1231	4.7170
A_3	4.1231	1.4142	4.7170

The computational complexity of the transportation problem is based on the simplex algorithm, which has an exponential worst case.

2.3. Image Retrieval Using the EMD.
Color distribution data applied for CBIR models is extracted from color segmentation images. As shown in Figure 3, the distribution consists of the weight, flow cost, and image feature. The weight is the number of pixels in each region. The flow cost of each distribution can be obtained by calculating the Euclidean distance between the image features in each region. The image feature of each region consists of the center position (x, y) and the average color signature (RGB). Figure 4 shows
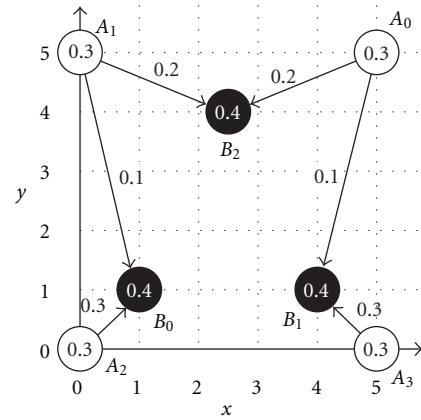

FIGURE 2: Optimized flow.

■ Weight
  ■ The number of pixels in each region
■ Flow cost
  ■ The Euclidean distance between image features
■ Image feature
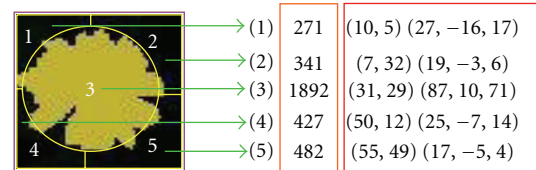 ■ The center position and color signature in each region

FIGURE 3: Color signature for the EMD.

a sample of retrieval results using the EMD, where this sample is retrieval results from 10,000 image dataset. On the other hand, Figure 5 shows a sample of retrieval results using the L_2 from the same dataset. From these results, it can
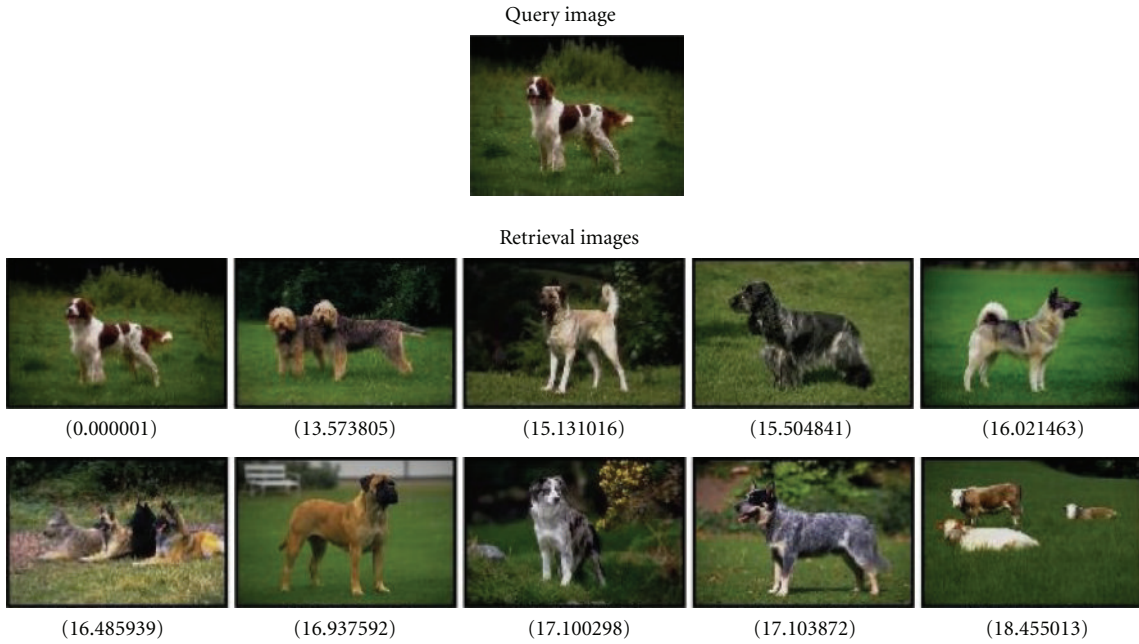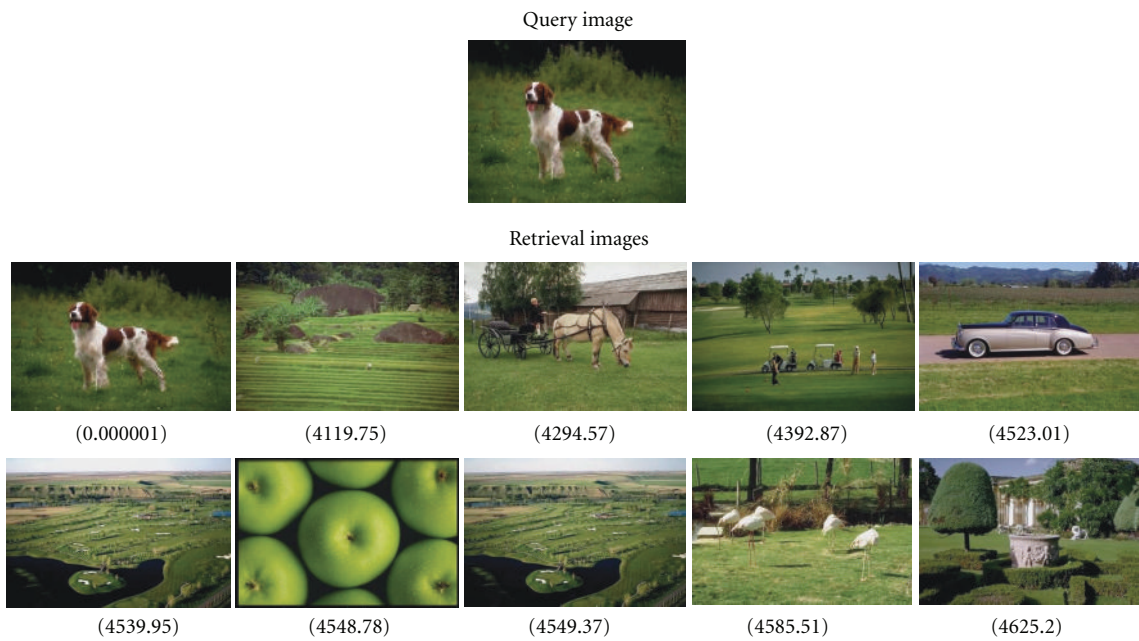
Query image

Retrieval images

(0.000001)   (13.573805)   (15.131016)   (15.504841)   (16.021463)

(16.485939)   (16.937592)   (17.100298)   (17.103872)   (18.455013)

FIGURE 4: Sample of retrieved results using the EMD.

Query image

Retrieval images

(0.000001)   (4119.75)   (4294.57)   (4392.87)   (4523.01)

(4539.95)   (4548.78)   (4549.37)   (4585.51)   (4625.2)

FIGURE 5: Sample of retrieved results using the L_p.

be found that CBIR systems using the EMD can reflect the human perceptual similarities.

3. EMD Lower Bound

3.1. Basic Definitions and Notations. We denote a signature x as follows:

$$x = \{(x_1, w_1), (x_2, w_2), \ldots, (x_n, w_n)\}$$
$$\equiv (X, w) \in D^{d,n}, \tag{9}$$

where

$$x = [x_1 \cdots x_n] \in R^{d \times n}, \quad w \geq 0. \tag{10}$$

Here, d is the dimension of the cluster $x_i \in R^d$, and n is the number of clusters. The total weight of the signature x is

$$w_\Sigma = \sum_{j=1}^{n} w_j. \tag{11}$$

In the case of the k-nearest neighbor search task, let r be the distance between the query image and the the kth candidate data during the search process.

3.2. Centroid-Based Lower Bounds.

The centroid $\overline{x}$ of the signature $x = (X, w) \in D^{d,n}$ is defined as

$$\overline{x} = \frac{\sum_{j=1}^{n} w_j x_j}{w_\Sigma}. \tag{12}$$

Theorem 1. *Suppose that* $x = (X, w) \in D^{d,m}$ *and* $y = (Y, u) \in D^{d,n}$ *are signatures of the equal total weight* $w_\Sigma = u_\Sigma$. *Then,*

$$EMD_{CB}(x, y) = ||\overline{x} - \overline{y}|| \le EMD(x, y). \tag{13}$$

Here, the ground distance $|| \cdot ||$ *is any* L_p *norm used to measure* $d(x_i, y_j)$, *as described in Cohen and Guibas [4].*

3.3. Calculation Skip Using the EMD_{CB}.

In the calculation of EMD_{CB}, the method described in Ajiok et al. [8] can be applied to skip its calculation steps, because EMD_{CB} is the L_2 distance between the centroid of each signature. This method regards the calculation of each centroid as each dimension in the sigunature vector.

An example is shown in which the signature A (Table 1) is the query, the distance between the signature C (Table 4) and the query is calculated, and r is set to 1.5. Then, the following centroids are obtained by (12):

$$\overline{A} = (2.5, 2.5),$$
$$\overline{C} = (4.5, 3.25). \tag{14}$$

First, the L_2 distance of the first dimensional centroid is calculated as below:

$$r^2 = 2.25 < (2.5 - 4.5)^2. \tag{15}$$

The distance of the first dimension is larger than r^2; hence, the calculation steps after the second dimension can be eliminated.

3.4. Independent Minimization Lower Bounds.

EMD_{CB} denotes the lower bounds based on the L_p distance. On the other hand, Assent et al. [5] have proposed independent minimization lower bounds based on cross-bin dissimilarity measures.

Theorem 2. *Suppose that* $x = (X, w) \in D^{d,m}$ *and* $y = (Y, u) \in D^{d,n}$ *are signatures of the equal total weight* $w_\Sigma = u_\Sigma$. *Then,*

$$EMD_{IM}(x, y) = \min\left(\sum_{i=1}^{m} \sum_{j=1}^{n} \frac{d_{ij}}{M} f_{ij}\right) \tag{16}$$
$$\le EMD(x, y).$$

TABLE 4: An example of the signature C.

	Weight	x	y
C_0	0.5	5	5
C_1	0.5	5	0
C_2	0.2	2	7

TABLE 5: Flow in EMD_{IM}.

| | w_{B_0} | w_{B_1} | w_{B_2} |
	0.4	0.4	0.4
w_{A_0} 0.3	0.0000	0.0000	0.3000
w_{A_1} 0.3	0.0000	0.0000	0.3000
w_{A_2} 0.3	0.3000	0.0000	0.0000
w_{A_3} 0.3	0.0000	0.3000	0.0000

Equation (16) *is obtained with the following constraints:*

$$f_{ij} \ge 0, \quad 1 \le i \le m, \ 1 \le j \le n, \tag{17}$$

$$\sum_{j=1}^{n} f_{ij} = w_i, \quad 1 \le i \le m, \tag{18}$$

$$f_{ij} \le u_j, \tag{19}$$

$$\sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij} = w_\Sigma = u_\Sigma, \tag{20}$$

where f_{ij} *is the flow from* x_i *to* y_j, $d_{ij} = d(x_i, y_j)$ *is the ground distance between cluster* x_i *and* y_j, *and* $M = \sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij} = w_\Sigma = u_\Sigma$.

3.5. Calculation Skip Using EMD_{IM}.

The calculation cost of the EMD_{IM} can be reduced by weakening the constraints of the transportation problem. In Figure 2, the total flow equals the total (supplies or demands) weight, because the constraints of (4) and (5) are satisfied. The difference between the EMD_{IM} and the normal EMD is the constraint on the amount of flows that one cluster may receive. While the normal EMD requires the sum of flows to a certain cluster to be equal to the weight of the cluster (18), the EMD_{IM} only ensures that for any single cluster, the incoming flows do not exceed the weight of the cluster (19); that is, the total flow to a cluster can exceed its weight. The final flows of $EMD_{IM}(A, B)$ are shown in Table 5 in which the total flow to B_2 exceeds w_{B_2}.

From this characteristic, the flows can be calculated for each supply cluster. To calculate EMD_{IM}, the method described in Ajiok et al. [8] can be applied to skip the calculation step of each supply cluster. This method regards the flow calculation of each supply cluster as each dimension in the signature vector.

An example of the calculation skip of the $EMD_{IM}(A, B)$ is shown, where r is set to 1.5. First, a minimum value of total transportation costs from A_0 is obtained, and its value is compared with r, as shown below:

$$0.3 \cdot 2.6926 = 0.80778 \tag{21}$$
$$< r.$$
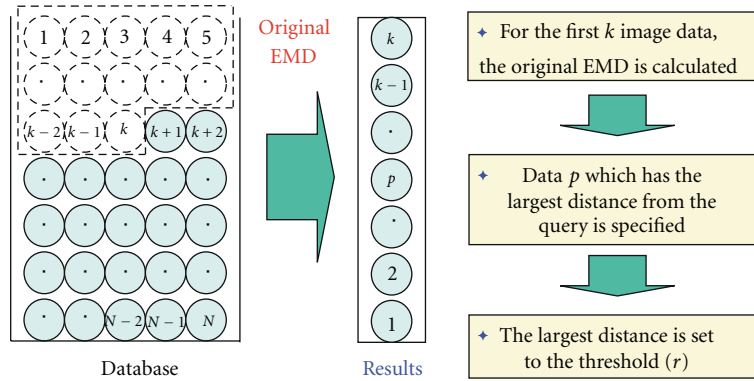
FIGURE 6: Algorithm for the first k image data.
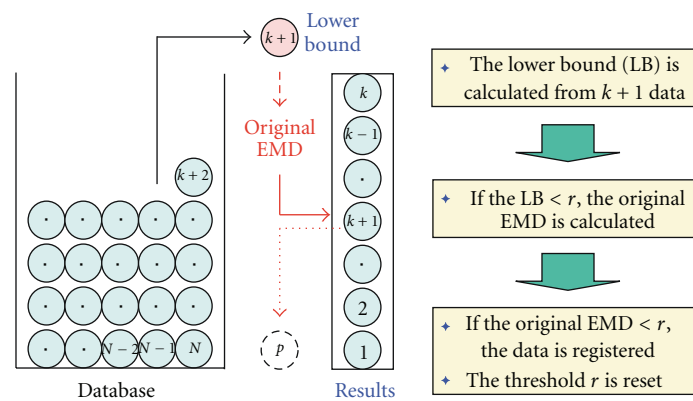
FIGURE 7: Algorithm from the $k + 1$ image data.

The value of A_0 is smaller than r; the total transportation costs from A_1 are added to A_0 as follows:

$$0.80778 + 0.3 \cdot 2.6926 = 1.61556$$
$$> r, \tag{22}$$

where the total value is larger than r; hence, the calculation of A_2 and A_3 can be eliminated.

4. Fast Retrieval Algorithm for the EMD

4.1. Algorithm Using the Lower Bounds. In the case of the k-nearest-neighbor search task, the retrieval algorithm for the first k image data is shown in Figure 6, and the retrieval algorithm from the $k+1$ image data is shown in Figure 7. This method calculates the original EMD for the first k images in the database, where the largest distance is set to the threshold r. This method calculates the lower bounding distance for succeeding data from $k+1$ image. This method does not calculate the original EMD if the lower bounding distance exceeds the threshold r, because the true EMD cannot be lower than the first k images. In all other cases, this method calculates the original EMD, and if the true EMD is lower than the first k images, a new image is registered, and the threshold r is reset.

4.2. Algorithm Using the Priority Queue. In a nearest-neighbor search based on a linear search, k search results are obtained when k features close to a given query are found. When a newly found result has a distance to the search feature that is smaller than the largest of the current k results, the results are updated. The data with the largest distance must be deleted from the candidate list, and the new result must be added. The efficient execution of this procedure is provided by using a data structure called a priority queue.

There are various implementations of the priority queue, one of them being *heap*. The heap structure is shown in Figure 8. The diagram shows an array as a binary tree, with a key specifying the two lower keys. Every key is larger than its two child keys. Such a data structure is called a complete binary tree; the heap is defined as a complete binary tree, which is represented as an array in which all the nodes satisfy the heap condition. Here, the heap condition is that the key in any node is larger than (or equal to) the keys in its child nodes. If this heap condition is satisfied, then the largest key is always in the root node. As shown in Figure 8, the root is numbered as 1, its children as 2 and 3, their children as 4, 5, 6, and 7, and so on. In such a representation, the parent of a node $a[j]$ is $a[j/2]$ (truncated at the decimal point if j is odd), and its children are $a[2j]$ and $a[2j + 1]$; thus, the parents and children can be identified easily for every node.
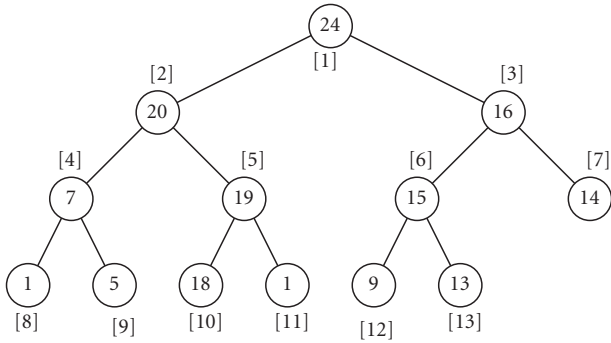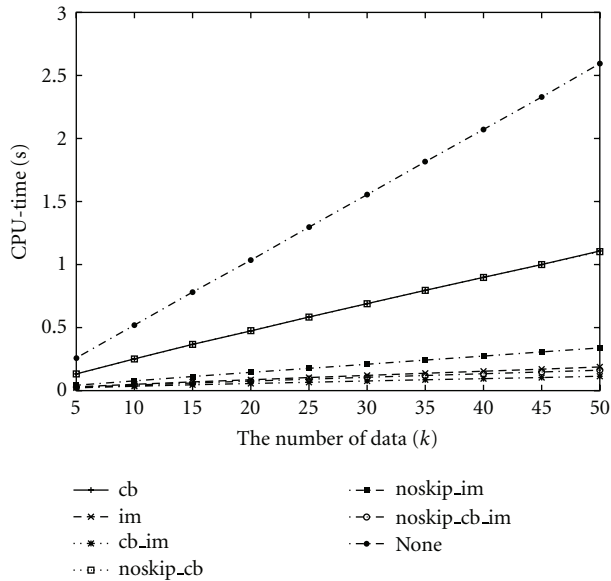
FIGURE 8: Heap structure.

FIGURE 9: CPU time on 5-dimensional data.

When a new search result offers a distance smaller than the largest value among the current results, the heap root (which has the largest value in the candidate list) is replaced with this newly found value. When this procedure causes a violation of the heap condition, the root value is exchanged with the largest of its children to restore the condition. If the heap condition is violated at the next level, a similar operation is applied; the procedure is terminated if a parent's value is smaller than that of both children, or when the heap bottom is reached. This processing ensures that the data structure can be repaired so that every node meets the heap condition.

5. Evaluations

5.1. Experimental Results. In order to evaluate this method, the number of color images in the database was increased from 5,000 to 50,000 (in increments of 5,000), and the retrieval time (CPU time) to search the top 10 images was evaluated. To evaluate the number of dimensions of the cluster representative x_i, we prepared 5- and 11-dimensional
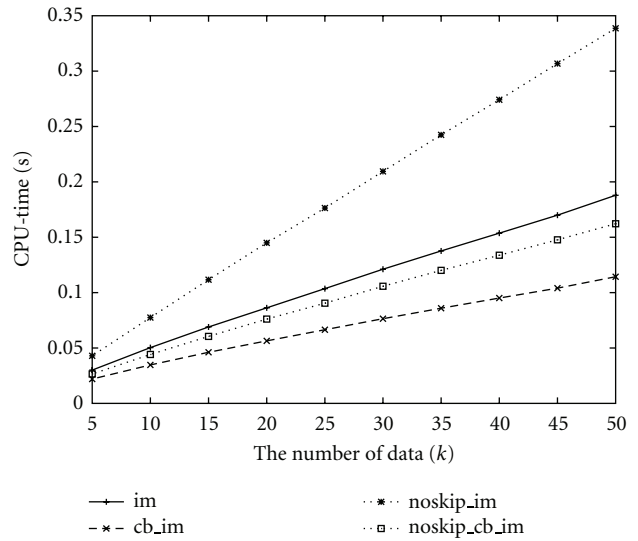

FIGURE 10: CPU time on 5-dimensional data (enlarged graph).

data as the image feature. These experiments were simulated using a PowerPC G4 running at 1.5 GHz with 1.25 GB of memory.

The experimental results are shown in Figures 9–12. The experimental results for the 5-dimensional data are shown in Figures 9 and 10, and the results for the 11-dimensional data are shown in Figures 11 and 12. In order to make the effect of the combination of the proposed methods comprehensible, Figures 10 and 12 are enlarged graphs of Figures 9 and 11 correspondingly. In these graphs, "none" indicates the normal EMD calculation, "cb" indicates the corresponding result obtained using the centroid-based lower bounding distance, "im" indicates the corresponding result using the independent minimization lower bounding distance, and "cb_im" indicates the corresponding result by combining the "cb" and "im" methods. In addition, "noskip_*" represents the value obtained by each method without the calculation-skipping algorithm.

From the experimental results, it can be seen that when the number of color images is 50,000 and has 5 dimensions, it takes approximately 2.5 seconds to retrieve the result for a query image by using the normal EMD calculation. The combination method of the lower bounds could obtain the same result as the normal EMD in approximately 0.15 second. On the other hand, the method using the calculation-skipping algorithm took only 0.1 second.

5.2. Related Works. Some derivation algorithms of the EMD have been proposed. Pele and Werman [9] proposed the fast algorithm of $\widehat{\text{EMD}}$, and Ling and Okada [10] proposed the fast algorithm of EMD_L_1. The $\widehat{\text{EMD}}$ can be obtain by replacing the ground distance of the EMD with the thresholded distance. EMD_L_1 is a replacement of the ground distance with the L_1 distance. Both of $\widehat{\text{EMD}}$ and EMD_L_1 are different from the original EMD. And they only indicate high-speed calculation methods in the case that the thresholded distance
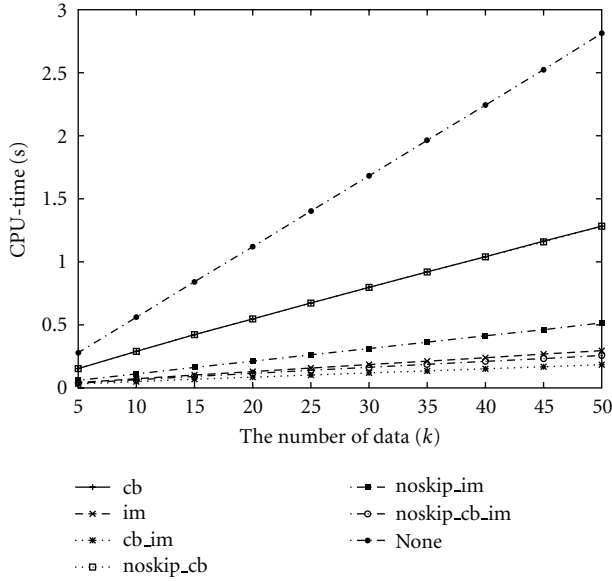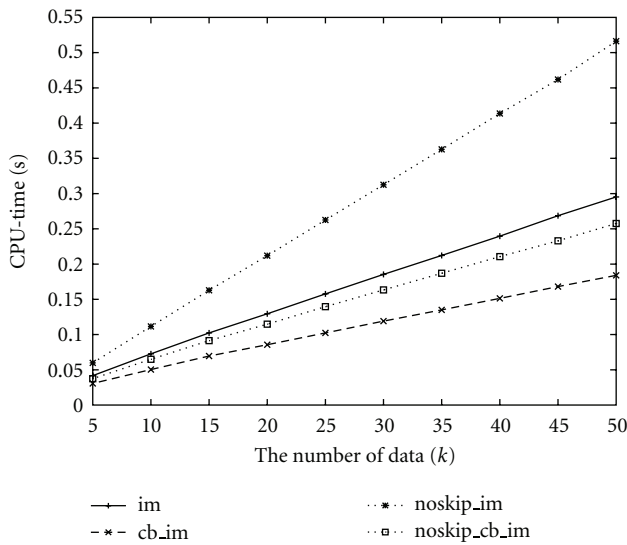
FIGURE 11: CPU time on 11-dimensional data.

FIGURE 12: CPU time on 11-dimensional data (enlarged graph).

and L_1 distance were used as the ground distance. On the other hand, by using our algorithm, the original EMD can be calculated efficiently. The characteristic of our algorithm is that unnecessary calculations can be skipped by using the EMD lower bounds, and the viewpoint of speedup is different from their methods.

5.3. Future Works. The proposed method selects first k image data in the database as the initial retrieval results (Figure 6). However, if there are some outliers in first k image data, the effect of speedup will become less, since the threshold r is much larger. Then, the smaller the distance between the query and first k image data is, the more the retrieval efficiency improves. In order to solve this problem, we are

TABLE 6: A preliminary experimental result.

	distance$_{\text{top}}$	distance$_{\text{vp}}$
query$_1$	5028.37	4723.46
query$_2$	5115.23	2264.38
query$_3$	4401.71	1969.05
query$_4$	4071.22	2437.28
query$_5$	4329.04	1588.82
average	4589.11	2596.60

planning to use the hierarchical clustering index to select the initial data. VP-tree [11] and M-tree [12], which are famous metric space indices, can be used as the hierarchical clustering index. By using these indices, appropriate initial data as first k image data can be obtained from the leaf node which is reached for the first time on the retrieval process. To be sure, these data are not true correct, but the possibility including outliers becomes low considerably. And it seems that the distance between the query and first k image data becomes small. The above improvement will be done as the future works.

In order to validate that the hierarchical indexing tree can help to reduce the effects of outliers in the first k images, the following preliminary experiment was done. First, the max distance between the query and top 100 image data is calculated. This distance is called distance$_{\text{top}}$ and set to the initial threshold r on the proposed method. Next, we obtain the max distance between the query and about 100 image data in the leaf node, which is reached for the first time by using the VP-tree. This distance is called by distance$_{\text{vp}}$, and set to the initial threshold r' on the improved method by using the VP-tree. These distances were calculated for 5 queries and the image database, in which was registered 10,000 image data. The preliminary experimental result is shown in Table 6, where the Euclid distance is used as the distance measure, and the number of dimensions is 48. From the preliminary experimental results, it is found that the initial threshold can be set to smaller value by using the VP-tree.

6. Conclusion

In this paper, a fast retrieval method for the earth mover's distance in multimedia databases is proposed. This method uses the lower bounding distance of EMD and combines it with a calculation-skipping algorithm. Moreover, the validity of the proposed method is supported by empirical observations. For future works, various multimedia information retrieval systems using the fast EMD should be implemented.

Acknowledgment

References

[1] C. Faloutsos, R. Barber, M. Flickner et al., "Efficient and effective querying by image content," *Journal of Intelligent Information Systems*, vol. 3, no. 3-4, pp. 231–262, 1994.

[2] Y. Rubner and C. Tomasi, *Perceptual Metrics for Image Databases Navigation*, Kluwer Academic Publishers, Dordrecht, The Ntherlands, 2001.

[3] M. Shishibori, Y. Ohnishi, S. Tsuge, and K. Kita, "Similar music retrieval for the query-by-humming using the Earth Mover's distance," *Transaction of Information Processing Society of Japan*, vol. 48, no. 1, pp. 300–311, 2007.

[4] S. Cohen and L. Guibas, "The Earth Mover's distance: lower bounds and invariance under translation," Tech. Rep. CS-TR-97-1597, Stanford University, 1997.

[5] I. Assent, A. Wenning, and T. Seidl, "Approximation techniques for indexing the earth mover's distance in multimedia databases," in *Proceedings of the 22nd International Conference on Data Engineering (ICDE '06)*, p. 11, usa, April 2006.

[6] W. Niblack, R. Barber, W. Equitz et al., "QBIC project: querying images by content, using color, texture, and shape," in *Storage and Retrieval for Image and Video Databases*, vol. 1908 of *Proceedings of SPIE*, pp. 173–187, San Jose, Calif, USA, February 1993.

[7] M. A. Stricker and M. Orengo, "Similarity of color images," in *Storage and Retrieval for Image and Video Databases III*, vol. 2420 of *Proceedings of SPIE*, pp. 381–392, San Jose, Calif, USA, February 1995.

[8] S. Ajiok, S. Tsuge, M. Shishibori, and K. Kita, "Fast multidimensional nearest neighbor search algorithm using priority queue," *IEEJ Transactions on Electronics, Information and Systems*, vol. 126, no. 3, pp. 353–360, 2006.

[9] O. Pele and M. Werman, "Fast and robust earth mover's distances," in *Proceedings of the 12th International Conference on Computer Vision (ICCV '09)*, pp. 460–467, October 2009.

[10] H. Ling and K. Okada, "An efficient earth mover's distance algorithm for robust histogram comparison," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 5, pp. 840–853, 2007.

[11] P. N. Yianilos, "Data structures and algorithms for nearest neighbor search in general metric spaces," in *Proceedings of the 4th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 311–321, January 1993.

[12] P. Ciaccia, M. Patella, and P. Zezula, "M-tree: an efficient access method for similarity search in metric spaces," in *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pp. 71–79, 1995.