

Optimal Keys for Image Database Indexing

Michael S. Lew D.P. (Nies) Huijsmans
Department of Computer Science
Leiden University, Postbus 9512
2300 RA Leiden, The Netherlands
{mlew huijsman}@wi.leidenuniv.nl

Dee Denteneer
Philips Research Laboratory
Prof. Holstlaan 4
5656 AA Eindhoven
dentenee@natlab.research.phil
philips.com

Topics: color image database indexing, query by content, content based image retrieval

Abstract. This paper examines the problem of efficient indexing of large image databases using the theory of optimal keys. The methods based on optimal keys are compared empirically with a texture classification method and template matching for benchmarking purposes in the Leiden WWW color image database and the 19th century portrait database. The different indexing methods are compared and evaluated in the problem space of finding copies of corrupted images. Real world noise is present in the form of print-scanner noise and general image degradation.

1 Introduction

Although the technology to digitize, transmit, and store images and video exists, we are still unable to analyze multimedia databases for semantically meaningful information. There have been a variety of different projects toward indexing multimedia databases (see Gudivada and Raghavan [1995] for an overview), but relatively little research has been focused toward comparing the different database indexing methods.

In order to compare the methods, we require an absolute ground truth. A useful application which as easily established ground truth is corrupted copy location. This would be indicative of a search for the original of a newspaper or magazine picture. Two interesting applications of copy location are *print-scanner noise* and *general image degradation*. Print-scanner noise is indicative of the performance of an algorithm for finding copies of scanned in newspaper or magazine pictures. An original image is printed via laser printer to plain paper and then scanned-in. For the general image degradation, the copies of the original image have been subjected to decades of different handling and treatment by their owners.

Most of the previous image database indexing methods have been based on features such as color histograms, moments, textures, etc. These are interesting methods, but none of them addresses the goal of the indexing, which is to locate

the true copy of a corrupted image. In this paper, we approached the problem from a database theory viewpoint using the theory of *keys*. A key can be as simple as the social security number of an individual, or more complex such as a hashing function. The important aspect of keys is that an *optimal key* is able to locate the desired item accurately and efficiently. In our formulation, an optimal key maximizes the probability that the key computed from a corrupted image is closer to the key computed from its original than to the key computed from a different image.

One basic indexing method is to perform sum of squared difference template matching. Another interesting and typical method is to use texture classification for finding similar images [Gudivada and Raghavan 1995]. Thus, we chose to compare the optimal key based methods to template matching and texture classification techniques.

In this paper, we evaluate optimal key algorithms for the purpose of noisy image copy location in the Leiden WWW color image database and the 19th century portrait database (LCPD). Section 2 describes the image databases and creation of the ground truth. Section 3 introduces and discusses optimal keys, local binary patterns, and template matching. In Section 4, we test the methods on the applications of print-scanner noise and general image degradation. Conclusions are given in Section 5.

2 Image Databases and Ground Truth

The Leiden WWW (World Wide Web) image database is an ongoing project to create an image index to the entire WWW. See the section called *WWW Demo Programs* for the http URLs of the databases. The size of the current online database is 100,000 images from the WWW.

The LCPD (see Figure 1) is currently composed of 5570 photographic images taken during the 19th century, and it will be continually expanded until at least 100,000 images are in the database. When the originals were taken, copies were typically created and distributed to different owners.



Figure 1. Sample images from the Leiden 19th Century Portrait Database.

In the print-scanner experiments, the original image is printed to plain paper using a laser printer and then scanned in. Examples from the grayscale LCPD database and the color WWW image database are shown in Figures 2 and 3.



Figure 2. Image 140 and print-scanner degraded copy

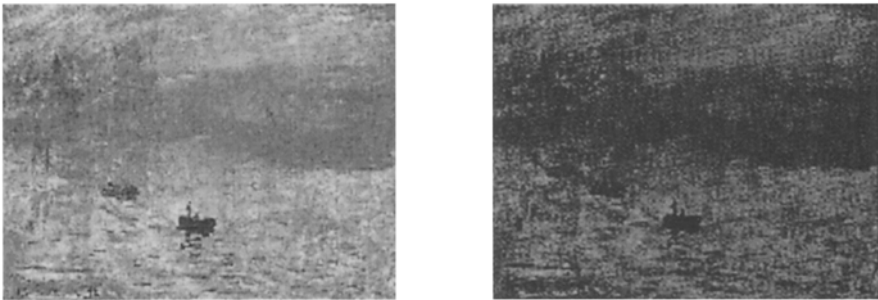


Figure 3. Monet's Sunrise from the Louvre WebMuseum and print-scanner copy

In the general image degradation experiments, the copies of the original were subjected to decades of general noise which includes scratches, writing, environmental exposure, etc. An example of the copies is shown in Figure 4. Note that these copies have different contrast and markings.

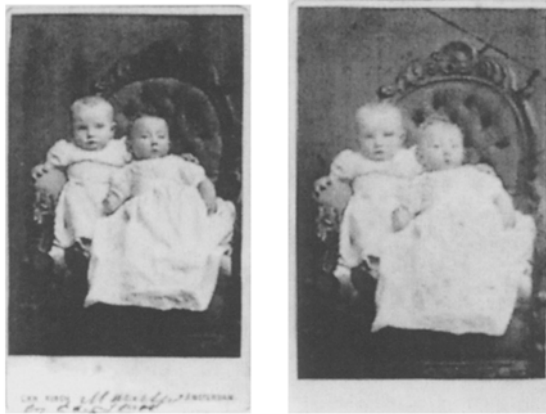


Figure 4. Two copies (c000409 and c000412) subjected to general noise.

3 Optimal Keys, Local Binary Patterns, and Template Matching

In this section, we discuss the three general indexing methods which we used in our experiments. These are (1) template matching, (2) texture classification, and (3) optimal key based methods. Our implementation of template matching is to minimize the sum of the squared differences (SSD) between two images, X and Y . The texture classification method of the local binary patterns (LBP) was chosen because it performed well in a recent survey [Ojala, et al. 1996]. Specifically, the LBP has the lowest single feature error rate for "Image Set 1." The effectiveness of optimal keys depends on how well the noise is modeled. Furthermore, for tractability, we make the assumption of linear keys in our derivation of the methods.

Keys are functions which map images onto one or more feature values. Instead of computing the SSD of two images (template matching), X and Y .

$$\sum_j \sum_i (X(i, j) - Y(i, j))^2 \quad (1a)$$

We need only compute the key difference, which is

$$\sum_m (k(X, m) - k(Y, m))^2 \quad (1b)$$

where

X and Y are two images with different content

m is the iterator over each key

Assuming that n represents the noise then the *optimal key* algorithm is that k such that

$$\Pr\{|k(X+n) - k(X)| < |k(X+n) - k(Y)|\} \quad (2)$$

is maximized. Optimal keys maximize the probability that the key computed from a corrupted image is closer to the key computed from its original than to the key computed from a different image. If we make the assumption of linear keys,

$k(X+n) = k(X) + k(n)$, and implement the magnitude of a stochastic variable by the magnitude of its variance, then if Σ_X and Σ_n are the covariance matrices of X and n , the criterion function is

$$\text{maximize } \frac{k^t \Sigma_X k}{k^t \Sigma_n k} \quad (3)$$

or

$$\text{maximize } k^t \Sigma_X k \quad (4)$$

$$\text{subject to } k^t \Sigma_n k = 1 \quad (5)$$

which bears similarity to Fisher's discriminant [1936]. If we ignore the noise process then equations (4) and (5) become

$$\text{maximize } k^t \Sigma_X k \quad (6)$$

$$\text{subject to } k^t k = 1 \quad (7)$$

which is the Hotelling transform [1933].

The representative texture classification method for the comparison is the local binary pattern. Local binary patterns were introduced by Wang and He [1990] and evaluated with respect to other texture methods by Ojala, et al. [1996]. In their recent survey, the local binary patterns had the lowest error rate for the first image set among a set of 12 different texture representations. In the surveyed LBP method [Ojala, et al. 1996], a texture unit is a two-level 3×3 pattern where the threshold is the center pixel. This gives 2^8 or 256 possible texture units. The texture spectrum is the distribution of the 256 patterns. In our copy location problem space, we calculate the texture spectrum for each image and rank them by the sum of squared differences.

For the WWW image database, we used two color spaces, namely, RGB and Luv (see Pratt 1991 for an overview of color spaces). The main difference between the RGB and Luv color spaces is that the Luv model separates the luminance into L, and places the chromaticity into uv. For the template method, we average the SSD from each dimension of the color space. In the optimal key methods, each pixel contributes three elements instead of 1 to the vector representing the image. The LBP method is generalized to color by finding the texture units for each color space dimension independently. This results in a texture spectrum of $256 \times 3 = 768$ patterns.

4 Experiments: Print-Scanner and General Image Degradation

In this section, we compare the methods based on optimal keys, local binary patterns, and templates for the applications of print-scanner noise and general image degradation. For the methods based on optimal keys, we have two choices, namely equations (4) and (5) or (6) and (7). The choice depends on whether it is tractable to model the noise of the given application. For the print-scanner application, the noise can be modeled empirically so we use equations (4) and (5). For the general image degradation application, we ignore the noise process and use equations (6) and (7).

4.1 Print-Scanner Noise

For LCPD database, 50 images were printed to plain paper using an HP LaserJet 5si/mx, and then scanned in using an HP ScanJet IICx. Table 1 shows the distribution of the rank of the print-scanner experiments. For the WWW image database, 100 images were printed using the Tektronix Phasor 240 color printer and then scanned using the HP ScanJet IICx. The column labeled *Best* refers to the probability that the copy in the database was in the first rank. The other columns show the probability that the copy in the database was found in positions 2 - 6, 7 - 16, and 17 - 36. Column *Worst* shows the probability that the database copy was not among the first 36.

Table 1. Distribution of Position of Print-Scanner Degraded LCPD Images

Method	Best	Second	Third	Fourth	Worst
Intensity Template	0.06	0.04	0.06	0.02	0.82
Gradient Template	0.64	0.1	0.04	0.02	0.2
LBP	0.68	0.12	0.04	0.06	0.1
Optimal Key	0.86	0.06	0.0	0.02	0.06

For the grayscale image database test, the optimal key method has the greatest probability of finding the copy in ranks 1, 1-6, 1-7, and 1-36. The results of applying the template, LBP, and optimal key methods using the RGB and Luv color spaces on the WWW image database are shown in Table 2.

Table 2. Distribution of Position of Print-Scanner Degraded Color Images

Method	Best	Second	Third	Fourth	Worst
Template - RGB	0.03	0.02	0.03	0.01	0.91
Template - Luv	0.04	0.02	0.05	0.02	0.87
LBP - RGB	0.39	0.10	0.06	0.04	0.41
LBP - Luv	0.53	0.14	0.01	0.03	0.29
Optimal Key - RGB	0.51	0.03	0.07	0.04	0.35
Optimal Key - Luv	0.79	0.07	0.05	0.01	0.08

For the WWW color image database, the optimal key method using the Luv color space had the greatest probability of finding the copy in the top rank. However, the LBP method using the Luv color space performs better than the optimal key algorithm using the RGB space. In general, the Luv color space gives better results than the RGB space. In this experiment, the optimal key method had higher performance relative to the other methods.

4.2 General Image Degradation

Since the image degradation is general, the noise could not be modeled accurately. In Table 3, the distribution of the rank error for the general degradation is shown. In addition, these tests were performed only on the grayscale LCPD database because a similar color database could not be found.

Table 3. Distribution of Position of General Degraded Image

Method	Best	Second	Third	Fourth	Worst
Intensity Template	0.42	0.02	0.06	0.04	0.46
Gradient Template	0.32	0.12	0.26	0.0	0.30
Local Binary Pattern	0.26	0.2	0.06	0.12	0.36
Optimal Key	0.34	0.14	0.12	0.22	0.18

These results are of particular interest because the template method has the greatest probability of an image being located as the top rank. However, the optimal key method has a greater probability of the copy being found in ranks 1-6, 1-17, or 1-36. One explanation for this is that this occurs because the noise was ignored in the optimal key method.

Regarding computational efficiency, Table 4 shows the number of features used by each method

Table 4. Number of Features Used by Each Method

Method	# of Features - color	# of Features - grayscale
Intensity Template	Variable on size	47250
Gradient Template	Variable on size	47250
LBP	768	256
Optimal Keys	40	25

Overall, the methods which were based on the optimal key selection not only had greater accuracy than the other methods, but also required less than 10% of the features as LBP.

5 Conclusions

In this paper we applied methods based on optimal keys, local binary patterns, and template matching to the problem space of noisy image copy location for a color and grayscale image databases. For images which have been printed and then scanned in, the optimal key method had the least ranking error for both grayscale and color image databases. In the case of the real noise from decades of exposure to different environments and scanner noise, the optimal key method had the least ranking error with respect to the LBP and template methods. Overall, the optimal key methods performed significantly better than the LBP and template methods when the noise could be modeled accurately, and slightly better in the case of general noise. Furthermore, it was consistently found that the Luv color space gives better results than the RGB space.

With respect to computational efficiency, it was shown that the optimal key methods are related to principal component techniques. Optimal key selection techniques require fewer features for similar accuracy, which results in less computation. In the experiments, the methods based on the optimal key selection techniques were shown to require less than 10% of the features relative to the LBP and template methods, while still achieving equivalent or better accuracy.

WWW Demo Programs

The color image database can be accessed via

<http://ind156b.wi.leidenuniv.nl:8086/yurimage.html>

The LCPD can be accessed via

<http://ind156b.wi.leidenuniv.nl:8086/intro.html>

<http://ind156b.wi.leidenuniv.nl:2000/>

Acknowledgments

This research was supported by a grant from Philips Research, and the Netherlands Computer Science Research Foundation.

References

- Faloutsos, C., R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, W. Equitz, "Efficient and Effective Querying by Image Content," *Journal of Intelligent Information Systems*, pp. 231-262, 1994.
- Fisher, R.A., "The Use of Multiple Measurements in Taxonomic Problems," *Annals of Eugenetics (London)* 7, pp. 700-725, 1936.
- Gudivada, V. N., and V. V. Raghavan, "Finding the Right Image, Content-Based Image Retrieval Systems," *Computer*, IEEE Computer Society, pp. 18-62, Sept. 1995.
- Hotelling, H., "Analysis of a Complex of Statistical Variables into Principal Components," *Journal of Educ. Psychol.*, vol. 24, pp. 417-441, 498-520, 1933.
- Ojala, T., M. Pietikainen and D. Harwood, "A Comparative Study of Texture Measures with Classification Based on Feature Distributions," vol. 29, no. 1, pp. 51-59, 1996.
- Pratt, W. K., **Digital Image Processing**, 2nd ed., John Wiley & Sons, 1991.
- Wang, L. and D. C. He, "Texture Classification Using Texture Spectrum," *Pattern Recognition* 23, pp. 905-910, 1990.