1-1-1997

Similarity measures for compressed image databases

P Sangassapaviriya
University of Wollongong

Philip Ogunbona
University of Wollongong, philipo@uow.edu.au

Follow this and additional works at: <https://ro.uow.edu.au/infopapers>

Part of the [Physical Sciences and Mathematics Commons](#)

Recommended Citation

Sangassapaviriya, P and Ogunbona, Philip: Similarity measures for compressed image databases 1997, 203-206.
<https://ro.uow.edu.au/infopapers/2144>

Similarity measures for compressed image databases

Abstract

For image database applications it is desirable that functions such as searching, browsing and partial recall be done without the need to totally decompress the image. This has the advantage of alleviating possible burden and degradation that the network may suffer. Edge images derived from wavelet-compressed images are considered as index that can be queried by example. Zernike moment invariants are used as descriptors for the index edge image and the query sketch image. The descriptions are compared for the purpose of database searching. The query images were allowed to undergo translation, rotation, scaling and some deformation. Simulation results gave 90% recognition rate.

Keywords

similarity, image, measures, databases, compressed

Disciplines

Physical Sciences and Mathematics

Publication Details

Sangassapaviriya, P. & Ogunbona, P. (1997). Similarity measures for compressed image databases. IEEE Region 10 Annual International Conference: Speech and Image Technologies for Computing Telecommunications (pp. 203-206). IEEE.

SIMILARITY MEASURES FOR COMPRESSED IMAGE DATABASES

P. Sangassapaviriya and P. O. Ogunbona

School of Electrical, Computer and Telecommunications Engineering

University of Wollongong

NSW 2522

ps08@st.elec.uow.edu.au, pogna@elec.uow.edu.au

ABSTRACT

For image database applications it is desirable that functions such as searching, browsing and partial recall be done without the need to totally decompress the image. This has the advantage of alleviating possible burden and degradation that the network may suffer. Edge images derived from wavelet-compressed images are considered as index that can be queried by example. Zemike moment invariants are used as descriptors for the index edge image and the query sketch image. The descriptions are compared for the purpose of database searching. The query images were allowed to undergo translation, rotation, scaling and some deformation. Simulation results gave 90% recognition rate.

1. INTRODUCTION

The popularity of the Internet and the vast amount of available multimedia data is bound to pose some problems for multimedia resource discovery. These problems will be exacerbated by the fact that a lot of users will access the network via a low bit-rate communication link for some time to come. Image is one of the resources that have a huge storage requirement in a multimedia database application; browsing a full resolution image over the network may degrade the total network response time. The search time of image databases depends largely on the file format used in the storage and the search methods employed. Generally, the storage mechanism employed should support image database functionalities such as partial image recall and incremental image recall [1] in addition to efficient compression. It is desirable that the image compression method used allows the aforementioned functionalities without a need for full image reconstruction; such a method is described in [2].

The method described in [2] encodes an image using the wavelet transform and uses an edge image derived from the detail sub-bands of the image decomposition as the index. With this approach there is no need for a separate index which might occupy a large storage space. Furthermore, this method allows the user to interact with the database using a "query by visual example" [3]. One would expect that this method of query along with other techniques based on salient features of the images would be used in a practical system.

Similarity measure is a key factor in any database application. This is more so for image databases where

the content bears varying degree of importance to the multitude of users that may want to use the images. Furthermore, some users may pose a query to the database without necessarily having a full knowledge of the composition of the image of interest. Thus, similarity rather than an exact match is a more suitable search criterion for image databases. The unstructured nature of image data also makes conventional database search methods inappropriate. Indeed the search should be human intuitive and allow for the inexact way in which humans pose queries.

In a given scenario, a user query may lack entire knowledge of the target image and the search engine needs to cope with a low resolution query with possibly some scale change, rotation, translation, and some degree of deformity. To search an image effectively, the search engine should be robust and capable of a partial similarity measure. This paper proposes a similarity measure based on moment invariants and defined over features of the edge image derived from [2]. The proposed method is invariant to scale, rotation, and translation.

Section 2 of the paper describes image feature based on moment invariant and Section 3 briefly reviews similarity measure paradigms. Section 4 presents the simulation results and discussion. Concluding remarks are given in Section 5.

2. IMAGE FEATURES

Features extracted from an image (or object) uniquely identify the image (or object) in a specific dimension. For instance, colour histogram is an image descriptor in colour space and spatial relationship of objects in an image is a representation of the image in the *spatial-relationship* domain. The importance of "objective" image features is easily seen in the inadequacy of annotations generated by a human annotator; there are no generally accepted predefined words for all the contextual occurrence of objects in images. Thus, image representation is better defined through image features

In the image database literature, some of the image features that have attracted attention include colour histogram [4][5], object shape [6][7][8], transform coefficient [9][10], texture [11], image edge [3], spatial relationship of objects in an image [12][13], moment invariant [14][15], and their combination [16]. These features are usually compared under a particular distance

metric; the closer distance, the more similar are the images (or objects) represented.

In the application described in this paper, an edge image is generated from a compressed representation of the image and used as the index. An appropriate feature should not only account for the nature of the edge image but must take into consideration the nature of possible queries. The user will be required to draw a sketch of some object in the target image. Moment invariants are a good choice in describing both the image index and the sketch presented by the user.

2.1. Moment

Moment is an important image descriptor used in pattern recognition [17][15] and computer vision [14]. In general, moment extracts features of an image with respect to a reference axis or frame as a numeric value and is defined as,

$$m_{pq} = \iint_{-\infty}^{\infty} x^p y^q f(x, y) dx dy \quad (1)$$

The centralized moments are given as,

$$\mu_{pq} = \iint_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy \quad (2)$$

where,

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}$$

For example, the center of mass of an image can be described by using the two first order moment. Hu [18] defined a set of seven algebraic moment invariants based on the centralized and normalized version of (1). Hu's algebraic functions are invariant under rotation, reflection, translation and scale change.

From definition, the regular moment has the form of projection onto monomials, $x^p y^q$, that constitute a complete but non-orthogonal basis set. This property accounts for the information loss, suppression and redundancy reported in [19]. Moments based on orthogonal basis functions have been introduced by Teague [20] to overcome problems associated with regular moments. Zernike moments are the projection of an image function onto a set of Zernike complex polynomials [21] and they form a complete orthogonal set over the interior of the unit circle. The Zernike moments of order n with repetition l for an image function $f(x, y)$ can be defined in terms of the regular moments as

$$A_{nl} = \frac{n+1}{\pi} \sum_{k=1}^n \sum_{j=0}^q \sum_{m=0}^l (-i)^m \binom{q}{l} \binom{l}{m} B_{nlk} \mu_{k-2j-l+m, 2j+l-m} \quad (3)$$

with $n-|l|$ selected as even and $|l| \leq n$

$$B_{nlk} = -1^{(n-k)/2} \frac{[(n+k)/2]!}{[(n-k)/2]! [(k+l)/2]! [(k-l)/2]!} \quad (4)$$

for $(n-k)$ is even

Because of orthogonal property of Zernike moments, image information is captured in a non-redundant manner by the various moment orders. For example, the zero-

order moment represents mean intensity of an image function whereas the first order moments are related to centre of mass of an image. With appropriate normalization Zernike moments enjoy invariance to scale, rotation, and translation

3. SIMILARITY MEASURE PARADIGMS

Numerical taxonomy, which allows the categorization of objects, is based on similarity functions. Similarity plays an important role in the theories of knowledge and behaviour [22]. The success of any feature set in discriminating or categorizing images depends to some extent on the use of an appropriate similarity function. Tversky [22] noted that it is not always appropriate to compute a metric distance between features, it might well be adequate to compare some features qualitatively. Similarity measures can be classified into 3 categories, namely, metric based models, set-theoretic based models and decision-theoretic based models. These will be briefly reviewed in the sequel.

3.1. Metric-based Models

A metric space is a set X with a distance function: d where $X^2 \rightarrow \mathbb{R}$ such that $\forall x, y, z \in X$. d satisfies the three axioms that define a metric [7]. The Minkowski r -metric, which generalizes a distance function, is defined as

$$d_r(x, y) = \left[\sum_{i=1}^n |x_i - y_i|^r \right]^{1/r} \quad r \geq 1 \quad (5)$$

where x and y are two points in an n -dimensional space and $i = 1, 2, \dots, n$. The well-known Euclidean distance is a special case when r is equal to 2. The other cases of interest are when r is equal to 1 (city-block distance) or ∞ . When r is equal to ∞ , the distance can be defined as

$$d_\infty(x, y) = \max |x_i - y_i| \quad (6)$$

The distance function can also be defined under a fuzzy set. Given two fuzzy sets, A and B , the distance between them is defined as,

$$d_r(A, B) = \left[\sum_{i=1}^n |\mu_A(x_i) - \mu_B(y_i)|^r \right]^{1/r} \quad r \geq 1$$

where $\mu_{(A)}, \mu_B$ are membership function.

Numerical features such as moment invariants lend themselves to comparison through these distance measures.

3.2. Set-theoretic based models

The set-theoretic model derived by Tversky are appropriate for feature comparison rather than distance computation and takes into consideration the fact that some features do not satisfy the symmetric property of metrics. For a set of features, A, B associated with images a, b the Tversky defined a parametric family of similarity functions known as the contrast model,

$$s(a, b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A) \quad (7)$$

for some $\theta, \alpha, \beta \geq 0$

It is clear that this measure is more suitable for qualitative features.

3.3. Decision-theoretic based models

The basis of these models is that a set of p attributes are defined for a set of n images. A decision matrix $X_{n \times p}$ is a zero-one matrix indicating the possession or otherwise of each attribute by each of the n images. For two images i, j , let n_{ij} , $n_{\bar{i}j}$ and q_{ij} represent the number of attributes they have (i) in common, (ii) missing and (iii) in disagreement respectively. Table 1 shows a list of commonly defined similarity measure based on these variables[23].

Table 1: Some commonly used decision-theoretic measures

ANBERG $S(i, j) = \frac{n_{ij}}{n_{ij} + 2q_{ij}}$	JACCARD $S(i, j) = \frac{n_{ij}}{n_{ij} + q_{ij}}$
KULCYNski $S(i, j) = \frac{n_{ij}}{q_{ij}}$	SOKAL-SNEATH $S(i, j) = (n_{ij} + n_{\bar{i}j}) / (n_{ij} + \frac{1}{2}q_{ij} + n_{\bar{i}j})$

These measures are obviously suitable for qualitative features. They however suffer from the drawback that the features must be defined a priori for both images. This may be the case in specialised image databases but will rarely be useful in a "free form" image database. It should be noted that the set-theoretic measure of Tversky does not assume that features are predefined.

This paper uses Euclidean distance as a means of measuring similarity between two Zernike moment based features.

4. SIMULATION RESULTS AND DISCUSSION

The performance of the proposed scheme is tested by generating a database of ten edge images. These images are artificially generated so as to remove the bias from noise. The images labeled Tm1 through Tm10 consist of edge images of both geometrical and arbitrary shapes. Sketches of a car and a house are also included. The query images are formed from affine transformed versions of the original images. For example, the edge image Tm1r (see Figure 1) is a 90° rotated version of Tm1. A free-hand drawing of Tm1 (see Tm1d in Figure 1) is included as a realistic query that has suffered some deformation.

4.1. Simulation 1

Regular moment invariant and Zernike moments of two images namely, tm1 and tm3, are computed and compared against variation of Tm1 in terms of scale, rotation, and some degree of deformity. Figure 1 shows the image Tm1, Tm3, and variation of Tm1. The results are compared in Table 2 where the suffixes s (scale), r (rotation), d (deformation) signify the transformations to which the image has been subjected and the suffix t signifies truncation.

Figure 1 shows all images used in the first simulation

From the table it is evident that Zernike moments possess better recognizing power than regular moment invariants for a threshold of about 1.0. Note that only seven features have been computed for both invariants so as to allow a fair comparison.

Table2: Euclidean distance between the regular and Zernike moments of Tm1 and transformed version.

Image	Distance	
	Regular moment	Zernike moment
Tm1	0.0000000	0.0000000
Tm1r	0.0908710	0.002936325
Tm1s	2.4481845	0.074282618
Tm1sr	16.5035253	0.116382903
Tm1t	4.1794616	0.556020464
Tm1d	15.6431139	0.817933406
Tm1dr	15.6431139	0.817061109
Tm1ds	16.0681505	0.779649897
Tm1dsr	16.0681505	0.779472024
Tm3	2.9406353	1.374051439

4.2. Simulation 2

This simulation is devised to test the performance of Zernike moment invariants in a database application. The ten edge images in the database were used; the query images were generated by transforming the images in the database. Three classes (P1, P2, P3) of query images corresponding to three sets of transformation on the images in the database are created. P1 consists of ten scaled images, P2 consists of 20 scaled and rotated (90°)

images and P3 consists of 30 scaled, and rotated (30° and 90°) images. No attempt has been made to ensure that the classes are disjoint. Zernike moment invariants up to order 12 (i.e. 47 features) were computed for each of the images in the query sets and compared to image database features. The rates of recognition are shown in Table 3.

Table 3: Rate of recognition of scaled and rotated query images from the database (P1: scaling; P2: scaling & rotation (90°); P3: scaling & rotation (30° & 90°).

	Transformation of the query image		
	P1	P2	P3
Rate	90%	90%	72%

The recognition rate for transformation P3 is the lowest of the three classes. The reason for this performance is likely due to the artifacts introduced by the rotation through 30°. This may suggest that the proposed method will be sensitive to noise. In [17] the problem of noise was anticipated and an averaged feature vector of a set of training images was used as the representation in the database. This method was reported to have given a recognition rate up to 99% for alphabetic characters; their worst case being 86%

5. CONCLUSION

A method of describing the edge image derived from compressed images is proposed. An artificially generated image database was used to test its performance and the results obtained are indicative of its potential. Further work needs to be done in order to use the edge images derived directly from the compressed images.

6. ACKNOWLEDGEMENT

This work is supported by the Australian Research Council through the ARC Small Grant Scheme.

7. REFERENCES

- [1] White, J. and A. Klinger. Image Coding in Visual Databases. in Proceedings of the IFIP TC2/WG2.6 Second Working Conference on Visual Database Systems. 1992. Budapest, Hungary: North Holland.
- [2] Ogunbona, P.O. and P. Sangassapaviriya. Content-based Retrieval from Compressed-Image Databases. in IEEE Data Compression Conference 1997. 1997. Snowbird, USA.
- [3] Hirata, K. and T. Kato. *Query by Visual Example: Content based Image Retrieval*. in *Advances in Database Technology EDBT'92, Third International Conference on Extending Database Technology*. 1992. Vienna, Austria.
- [4] Swain, M.J. *Interactive indexing into image databases*. in *SPIE*. 1993. San Jose, California.
- [5] Yihong, G., Z. Hongjiang, and C.H. Chuan. *An Image Database System With Fast Image Indexing Capability Based on Colour Histograms*. in *1994 IEEE Region 10's Ninth Annual International Conference*. 1994. Singapore.
- [6] Mumford, D., *Mathematical theories of shape: Do they model perception?* Geometric Methods in Computer Vision, 1991. 1570 *SPIE*: p. 2-10.
- [7] Pratt, W.K., *Digital Image Processing*. second ed. 1991, New York: John Wiley and Sons.
- [8] Mumford, D. *The problem with robust shape descriptions*. in *First International Conference on Computer Vision*. 1987. London, England.
- [9] Jacobs, E.C., A. Finkelstien, and H.D. Salesin, *Fast Multiresolution Image Querying*. 1995.
- [10] Pentland, A., R.W. Picard, and S. Sclaroff, *Photobook: Content-Based Manipulation of Image Databases*, in *Multimedia Tools And Applications*, B. Furt, Editor. 1996, Kluwer Academic Publishers: Boston. p. 42-80.
- [11] Equitz, W., *Using texture for query by image content (qbic)*. 1993, IBM Almaden Research Centre, San Jose, CA.
- [12] Rabitti, F. and P. Savino, *Efficient Content Based Retrieval in Image Databases: A Probabilistic Approach*. *Computer Vision for Industry*, 1994. *SPIE* vol.2185: p. 48-58.
- [13] Petrakis, E.G.M. and S.C. Orphanoudakis, *Methodology for the representation, indexing and retrieval of images by content*. *Image and Vision Computing*, 1993. 11(8): p. 457-520, October 1993.
- [14] Ghosal, S. and R. Mehrotra, *A Moment-Based Unified Approach to Image Feature Detection*. *IEEE Transactions on Image Processing*, 1997. 6(6): p. 781-793.
- [15] Dudani, S.A., K.J. Breeding, and R.B. McGhee, *Aircraft Identification by Moment Invariants*. *IEEE Transactions on Computers*, vol.C-26, no.1, p. 39-45, January 1977
- [16] Niblack, W., et al., *The QBIC Project: Querying Images By Content Using Color, Texture, and Shape*. *Computer Vision for Industry*, 1993. *SPIE* vol.1908: p. 173-181.
- [17] Khotanzad, A. and Y.H. Hong, *Invariant Image Recognition by Zernike Moments*. *IEEE Transaction of Pattern Analysis and Machine Intelligence*, vol.12, no.5, p. 489-497, May 1990.
- [18] Hu, M.-K., *Visual Pattern Recognition by Moment Invariants*. *IRE Transactions on Information Theory*, 1962. IT-8: p. 179-187.
- [19] Abu-Mostafa, Y.S. and D. Psaltis, *Recognitive Aspects of Moment Invariants*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.PAMI-6, p. 698-706, November 1984.
- [20] Teague, M.R., *Image analysis via general theory of moments*. *J. Opt. Soc. Am.*, 1980. 70(8): p. 920-930.
- [21] Born, M. and E. Wolf, *The Circle Polynomials of Zernike*, in *Principles of Optics*. 1980, Pergamon Press: Sydney. p. 767-774.
- [22] Tversky, A., *Features of Similarity*. *Psychol. Rev.*, 1977. 84: p. 327-352.
- [23] Joly, S. and G. Le Calve, *Similarity functions*, in *Classification and Dissimilarity Analysis*, Bernard Van Custem, Editor. 1994, Springer-Verlag: New York. p. 67-86.