TOWARD ROBUST MOMENT INVARIANTS FOR IMAGE REGISTRATION

Nawaf I. Almoosa, Soo Hyun Bae, and Biing-Hwang Juang

School of Electrical and Computer Engineering
Georgia Institute of Technology, Atlanta, GA 30332
{nawaf,soohyun,juang}@ece.gatech.edu

ABSTRACT

We apply pattern recognition techniques to enhance the robustness of moment-invariants-based image classifiers. Moment invariants exhibit variations under transformations that do not preserve the original image function, such as geometrical transformations involving interpolation. Such variations degrade the performance of classifiers due to the errors in the nearest neighbor search stage. We propose the use of Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA) to alleviate the variations and enhance the robustness of classification. We demonstrate the improved performance in image registration applications under spatial scaling and rotation transformations.

Index Terms— Moment invariants, principle component analysis, linear discriminant analysis, image registration

1. INTRODUCTION

Moment invariants (MIs) are a class of image descriptors first derived by Hu [1]. The general interest in MIs stems from their ability to describe the shape and photometric intensity of an image, and to remain invariant to several classes of degradations, e.g. affine transformations [2] and blurring [3]. Their compact representation reduces storage requirements, which is desired specially in colored image recognition. However, MI-based recognition deteriorates in the presence of common sources of image variations, which can be introduced at various stages such as capturing (e.g. camera noise), digitization (spatial and intensity quantization noise), and digital processing (e.g. image-resampling). The effects of these variations were analyzed in several works. For instance, the effects of noise and spatial quantization on some moment invariants were analyzed in [4, 5] respectively. However, no solution attempts were offered.

In this paper, we propose the use of techniques that can be flexibly applied to improve the robustness of moment invariants to several degradation sources. Our MI-based classification system employs linear discriminant analysis (LDA), and principal component analysis (PCA) to increase the recognition robustness. LDA is driven by observations generated by degrading several instances of each class and calculating their moment invariants. LDA is then applied to increase class clustering and separation. Then, PCA is applied to reduce the dimensionality of the feature space, and therefore alleviating redundancies and aiding with data visualization.

The organization of the paper is as follows: Section 2 outlines the mathematical theory of moment invariants and discusses their practical implications. Section 3 introduces the proposed technique. Section 4 demonstrates the improved performance of recognition in the chosen application of image registration. Finally is the conclusion in Section 5.

2. MOMENT INVARIANT DESCRIPTORS

MI-based image recognition started with Hu's seminary work [1], who derived MIs invariant to translation, scaling and rotation. Several works followed to analyze their noise performance [4], and derive invariants to new transformations, such as blurring [3], and as affine transformations [2].

To derive moment invariants, digital images are assumed to be continuous two-dimensional probability distribution functions. This enables describing images using common statistical quantities, and mapping them to achieve invariance. The statistical quantity of interest here is the central moment (CM) of order $(p+q)$, which given an image $f(x, y)$, can be defined as

$$C_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-x_c)^p(y-y_c)^q f(x,y)\,dxdy, \tag{1}$$

where $x_c = m_{10}/m_{00}$ and $y_c = m_{01}/m_{00}$ are the centriods of x and y respectively. The centroid subtraction renders the CMs translation-invariant. Furthermore, the CMs can be made scale invariant by normalization, yielding the *normalized central moments*, given by $\mu_{pq} = C_{pq}/C_{00}^{w}$ where $w = (p+q+2)/2$. Hu's MIs are mappings of second- and third-order normalized central moments, and they are rotation-, scale- and translation-invariant. Rigorous definitions of Hu's MIs are available in [1]. Subsequently, several MIs were derived, of which we pick the Affine Moment Invariants (AMI) [2]. AMIs are invariant to affine transformations, which map the coordinates (x, y) of an image to $(\hat{x}, \hat{y})$ according to

$$\hat{x} = a_{11}x + a_{12}y + a_{13}, \quad \hat{y} = a_{21}x + a_{22}y + a_{23}. \tag{2}$$

AMIs address the redundancy issues linked with Hu's moment invariants, and they can be derived up to any order. AMIs are used in the experiments in Section 4, and the reader is referred to [2] for their rigorous definitions and derivations.

In proving the invariance of a certain MI, the image and the transformation to which the MI is invariant are assumed to be continuous and noise-free. Such conditions do not exist practically, as images are represented by finite-precision pixels in a discrete coordinate system, and noise is introduced to images due to capturing devices and quantization effects. Therefore, several captured instances of the same scene are likely to exhibit variations in the

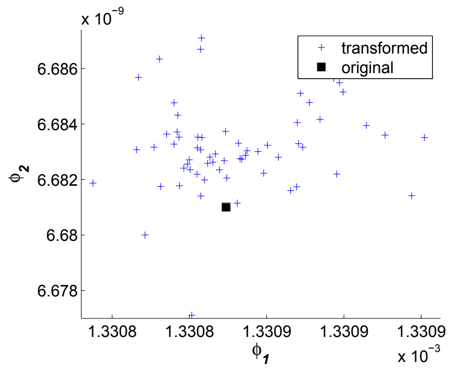


Fig. 1. Moment Invariants scatter plot for the first two AMIs of rotated and scaled instances. The image is “Lena”.

calculated moment invariants, which will vary based on the capturing and quantization noise levels. Teh *et al.* [4] analyzed the effects of noise on moment invariants, and have found that variations are inevitably introduced, and that higher order moments are more sensitive to noise. As for the effects of quantization, Salama *et al.* [5] have analyzed the disparity between ideal and quantized MIs calculated for square images, and have found that the error decreases as the image size increases, and the manner of decreasing may or may not be monotonic.

The restriction to the continuous domain also applies to transformations. MIs that are professed to be invariant to a certain transformation may not be so if the transformation is applied in the digital domain. For instance, several MIs were derived to be invariant under coordinate mapping transformations, e.g. rotations and affine transformations. These transformations may occur due to changes in the capturing device orientation, or are digitally induced. In the former case, the transformation is continuous and it is expected that moment invariants will exhibit minor variations (mainly due to quantization and capturing noise). In the latter case, variations are expected to increase due to the nature of digital transformations, which usually involve nonlinear operations (e.g. image-resampling which involves truncations of pixel values and rounding of coordinates). An example of the variations caused by such nonlinear effects is shown in Figure 1, which shows the space formed by the first two AMI moments calculated of rotated and scaled instances of the image “Lena”. [5] reported variations due to digitally-induced projection and rotation transforms respectively. Figure 1 also reveals a common difficulty with the calculation of MIs, namely the required precision. As higher moments are used, they are more aggressively normalized, leading to extremely minute numbers that require high precision.

3. PROPOSED SYSTEM

3.1. Classification System Definition

Our main contribution is the introduction of a data-driven class separation/clustering and dimensionality reduction to MI-based image recognition. Our approach performs the classification in a space with maximum clustering of classes, which reduces the likelihood of errors in the nearest neighbor search (NNS) stage. Also, we would like for this space to be of minimum size by eliminating the redundancies which may exist between MI elements. We design a combined algorithm using LDA for the first aim and PCA for the second.

Based on the expected degradation model, the class occupancy in the feature space is grown from a single point, which is the usual case in MI-based recognition, to multiple points corresponding to different degraded class instances. Therefore, decreased separation, and in some cases overlapping, between the regions occupied by classes is likely to occur. We use LDA to improve class discrimination by projecting the data to a space with maximized separation of classes. To aid with data visualization, and to eliminate redundancies that might exist between MI elements [2], the dimensionality of the LDA space is then reduced by calculating its principal components [6]. We now proceed with the rigorous definition of the aforementioned process.

Suppose an n-class system is given by $\beta = [b_1\, b_2\, \cdots\, b_n]^T$ where b_i is the i^{th} class, $i \in [1, n]$. We define the MI_m operator, which calculates the moment invariants of an image function $f(x, y)$, resulting in an m-element vector, where m is a function of the MI type and the order of the moments used.

$$\mathrm{MI}_m(f(x,y)) = \{\phi_1,\, \phi_2,\, \cdots\, \phi_m\} = \hat{\phi} \in \mathbb{R}^m. \tag{3}$$

If moment invariants are used as feature vectors for the classification system β, the class matrix $\alpha \in \mathbb{R}^{n\times m}$ is defined as $\alpha = [\alpha_1 \alpha_2 \cdots \alpha_n]^T$, where $\alpha_i = \mathrm{MI}_m(b_i)$, $\alpha_i \in \mathbb{R}^m$. We assume an expected degradation model $\mathbb{D} = [d_1\, d_2\, \cdots\, d_\psi]^T$ which consists of ψ types. Class observations are generated by applying the degradations in the set above to the objects in the class-system β. Each degradation is applied with its own set of parameters. The cardinality of each degradation parameter set is denoted as C_i, and therefore, the total number of observational instances per class is given by $L = \sum_{i=1}^{\psi} C_i$. After applying the degradations to the class-system β, and calculating their moment invariants, the observation matrix $A \in \mathbb{R}^{z\times m}$ (where $z = L \times n$) is formed as follows

$$A = [\hat{\alpha_1}\, \hat{\alpha_2}\, \cdots\, \hat{\alpha_n}]^T, \tag{4}$$

where $\hat{\alpha_i} \in \mathbb{R}^{L\times m}$. Each row of A represents an observation (degraded class instance), whereas each column belongs to a variable (i.e. an MI element). The observational-data generation process is illustrated in Figure 2.

3.2. Clustering and Dimensionality Reduction

LDA is used to project the data onto a space that maximizes the separation between classes while keeping classes as compact as possible. Let μ_i and S_i be the mean vector and covariance matrix of class $\hat{\alpha_i}$, the within-class scatter matrix S_w and the between-class scatter matrix S_b are

$$S_w = \sum_{i=1}^{n} L S_i, \quad S_b = \sum_{i=1}^{n} \mu_i \mu_i^T. \tag{5}$$

The fisher criterion [6] is given by

$$\frac{v_{lda}^T S_b v_{lda}}{v_{lda}^T S_w v_{lda}}. \tag{6}$$

Finding the projection vector v_{lda} that maximizes the above criterion ensures both the clustering of class members and increased separation between classes. The fisher criterion is maximized when the projection vector v_{lda} corresponds to the eigenvalues of the

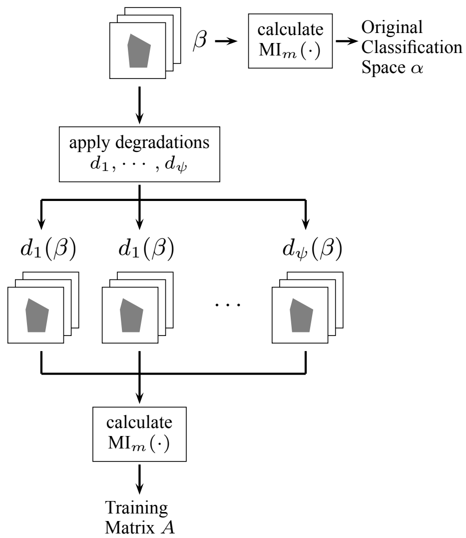
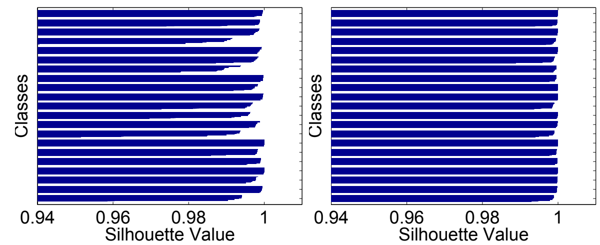
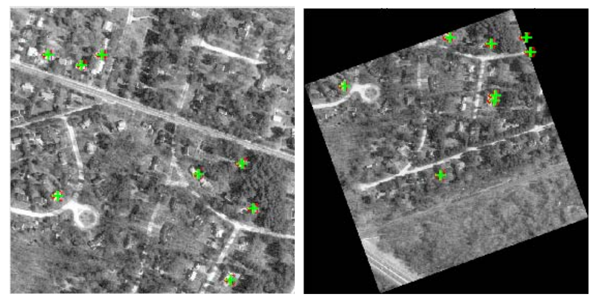


Fig. 2. A schematic of the observational data generation process of MIs for ψ-types degradation models, described in Section 3.1.



(a) AMI (b) AMI with LDA

Fig. 3. Silhouette plots for the original AMIs and the trained AMIs by LDA. A value of 1 corresponds to wide separation and -1 does to the complete overlap with other clusters.



(a) Reference image (b) Test image

Fig. 4. Example of POI-matching. The test image is translated and rotated instance of the reference image.

matrix $S_b^{-1} S_w$. Once the eigenvectors are calculated, they are used to project the observations matrix to a new one $A_{lda} = A \cdot v_{lda}$. After generating the space A_{lda}, a further step is needed to facilitate visual evaluation of the results and redundancy alleviation. A_{lda} is m-dimensional, and some dimensions may contain redundancy [2], therefore, the PCA is used to reduce the dimensionality of the feature space. Assuming W is the covariance matrix of A_{lda}, v_{pca} and λ are the eigenvectors and eigenvalues of W. The principal components can be found by sorting the eigenvectors in descending order of the corresponding eigenvalues. The p first columns of the sorted eigenvector matrix $\hat{v_{pca}}$ are used to project A_{lda} onto reduced dimensionality space $A_{pca} \in \mathbb{R}^{z \times p}$. The columns of A_{pca} are linear combinations of the columns of A_{lda} that maximize the variance, and hence they are denoted as the principal components.

Using v_{lda} and v_{pca}, the original MI feature space given by alpha can now be transformed into a new space, given by

$$\gamma = \alpha \cdot v_{lda} \cdot v_{pca}. \tag{7}$$

To classify a test case ξ, its moment invariants are projected in the same manner α is projected in (7), and then its nearest neighbor is found using a distance measure.

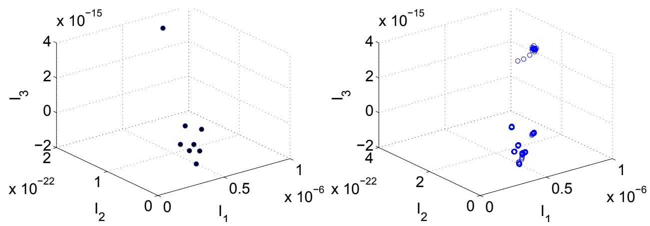
4. EXPERIMENTAL RESULTS

We present two experiments to demonstrate the performance improvements of the proposed system. In the first experiment, we formed a class system of 20 grayscale objects chosen from the Amsterdam Library of Object Images database [7], and applied a degradation model consisting of rotation and spatial scaling supported by bicubic-interpolation, and calculated their AMIs. Then, using LDA, the AMIs are projected onto a space with maximum separation of classes. We use the silhouette plot to evaluate how well the data is clustered in each space. The silhouette plot quantifies the closeness of a point in certain class to the points in the neighboring classes by assigning a value in the interval [-1, 1], where a value of 1 corresponds to wide separation and -1 corresponds to an overlap with neighboring clusters. The silhouette plots before and after LDA are shown in Figure 3 (a) and (b) respectively. Clearly, Figure 3 (b) has more values closer to 1, indicating better separation of classes.

The second testing benchmark for the proposed system is image registration, which consists of two main operations: determining if a group of images are representatives of the same scene, and combining them from multiple coordinate systems into a single one. The first task involves finding points of interest (POIs) in a test image and matching them to the POIs in a reference image. The POIs are unique corners that can be detected using specialized algorithms, here the Harris corner detector [8], and a correct registration requires correct matching between the reference and test POIs.

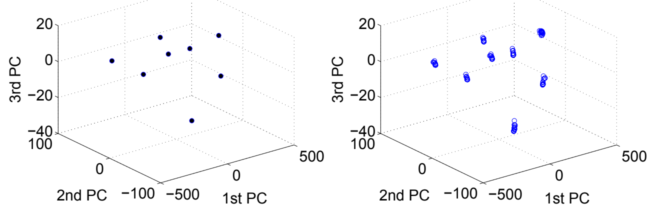
We compare our method with MI-based image registration system proposed in [3]. Each POI in the reference and test image is represented as the MIs of the circular region of which the POI is a center. POI-matching is then performed by finding reference and test POI pair that minimize a distance measure. The experiment consists of trying to match the test image shown in Figure 4 (b) to the reference image shown in Figure 4 (a). The test image is the lower third section of the reference image, degraded with a 20-degrees rotation employing bi-cubic interpolation. An attempt to register the test image using a radius of seven pixels and AMIs is shown in Figure 7. Correct registration is indicated by a black region in the difference image, corresponding to zero difference between the reference and the registered test image. Clearly, this is not the case with Figure 7 (a).

The incorrect registration is due to a false matching between reference and test POIs, which might be attributed to the poor sep-



(a) Original (b) Transformed

Fig. 5. POI Feature space.



(a) Original (b) Transformed

Fig. 6. POI Feature Space with the processing of LDA/PCA.

(a) AMI (b) AMI with LDA/PCA

Fig. 7. Examples of image registrations by the MI feature descriptors.

aration between POIs, as shown in Figure 5 (a). Applying a series of rotations and spatial scaling degradations to each class results in the feature space shown in Figure 5 (b), which validates the aforementioned point, since decreased separation, and in some cases overlap has occurred between some of the POIs. The dimensions in both figures are the first three AMI moments.

We consider the aforementioned rotation and scaling transformations as training data, and Figure 5 (b) represents the first three dimensions of the observation space A. We calculated the LDA and the PCA projection vectors, and projected the original POI feature space onto a new space. Using the first three principal components, the new POI feature space is shown in Figure 6 (a)-(b), where it is clear that the POI separation and clustering are significantly improved. Performing the matching between the reference and test POIs in the new space resulted in correct registration as indicated by the black region shown in Figure 7 (b).

Table 1 summarizes the results of POI matching using the system in [3] and the proposed system. Both systems were tested by applying several distorted images, where the distortions involve bi-cubic rotations and spatial scaling, and noting the percentage of times the systems perform the POI matching correctly. As the table shows, there is a significant improvement in the rate of suc-

Table 1. POI Matching Comparison

	System in [3]	Proposed
Avg. POI recognition	58.7%	87.6%
Max. POI recognition	70%	95%
Min. POI recognition	40%	75%
Avg. Pixel Displacement	0.2583	0.2305

cessful POI matching when the proposed system is used. In addition to that, the pixel displacement, calculated as the difference between the coordinates of the reference and matched POI is reduced, increasing the alignment of the reference and test images when combined to a single coordinate system.

5. CONCLUSIONS

We have presented a method for improving the robustness of moment-invariants-based recognition systems. The robustness of the classification system is improved by subjecting the system to expected degradations to generate observations which will guide the linear discriminant analysis to maximize the separation between classes. Principal component analysis is then applied to minimize the redundancy and aid with data visualization. Performance improvement is demonstrated with the application of image registration under the degradation caused by digital rotation and scaling. The results show that the trained system has better class discrimination, and is able to correctly register images that the untrained system is not able to. The proposed method can be applied in other image recognition applications with various degradation sources.

6. REFERENCES

[1] M. K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Information Theory*, vol. 8, pp. 179–187, 1962.

[2] J. Flusser and T. Suk, "Pattern recognition by affine moment invariants," *Pattern Recognition*, vol. 26, no. 1, pp. 167–174, 1993.

[3] J. Flusser, B. Zitova, and T. Suk, "Invariant-based registration of rotated and blurred images," in *IEEE Proc. Int. Geoscience and Remote Sensing Symp.*, 1999, vol. 2, pp. 1262–1264.

[4] C. H. Teh and R. T. Chin, "On image analysis by the methods of moments," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 10, no. 4, pp. 496–513, 1988.

[5] G. I. Salama and A. L. Abbott, "Moment invariants and quantization effects," in *IEEE Proc. Comp. Vision and Pattern Recog.*, 1998, pp. 157–163.

[6] Y. Koren and L. Carmel, "Robust linear dimensionality reduction," *IEEE Trans. Visual. Comput. Graphics*, vol. 10, no. 4, pp. 459–470, 2004.

[7] J. M. Geusebroek, G. J. Burghouts, and A. W. M. Smeulders, "The amsterdam library of object images," *Int. J. Comput. Vision*, vol. 61, no. 1, pp. 103–112, 2005.

[8] B. Zitova and J. Flusser, "Image registration methods: A survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, 2003.