Evolutionary Multi-Objective Optimization of Trace Transform for Invariant Feature Extraction

Wissam A. Albukhanajer*, Yaochu Jin*, Johann A. Briffa*

*Department of Computing,

University of Surrey,

Guildford, Surrey, GU2 7XH, UK

{w.albukhanajer, yaochu.jin, j.briffa}@surrey.ac.uk

Godfried Williams†

†Intellas UK Ltd,

Level 37, One Canada Square,

London, E14 5AA, UK

g.williams@intellas.co.uk

Abstract—Trace transform is one representation of images that uses different functionals applied on the image function. When the functional is integral, it becomes identical to the well-known Radon transform, which is a useful tool in computed tomography medical imaging. The key question in Trace transform is to select the best combination of the Trace functionals to produce the optimal triple feature, which is a challenging task. In this paper, we adopt a multi-objective evolutionary algorithm adapted from the elitist non-dominated sorting genetic algorithm (NSGA-II), an evolutionary algorithm that has shown to be very efficient for multi-objective optimization, to select the best functionals as well as the optimal number of projections used in Trace transform to achieve invariant image identification. This is achieved by minimizing the within-class variance and maximizing the between-class variance. To enhance the computational efficiency, the Trace parameters are calculated offline and stored, which are then used to calculate the triple features in the evolutionary optimization. The proposed Evolutionary Trace Transform (ETT) is empirically evaluated on various images from fish database. It is shown that the proposed algorithm is very promising in that it is computationally efficient and considerably outperforms existing methods in literatures.

Keywords—Trace transform; image recognition; invariant feature extraction; copyright protection; evolutionary algorithms; multi-objective optimization.

I. INTRODUCTION

A. Background

With the rapid growth of the Internet and increasing availability of editing software, intellectual property and copyrighted visual materials have become more vulnerable to a range of threats and infringements such as illegal copying, editing, reproduction and distribution. Digital watermarking is a good solution to the image authentication and copyright protection for digital images, which is accomplished by embedding watermarks into the digital image and subsequently using them to authenticate the image [1]. However, digital watermarking techniques are inappropriate for copyright enforcement of unauthorized images of cultural heritage artifacts, such as statues and paintings. For instance, digital images of such artifacts taken by users with a consumer camera and subsequently published on the Internet cannot be detected by watermarking techniques, as no watermark exists. Therefore, copyright protection of such images is

considered as an image identification problem instead of watermarking.

Image identification is of a great interest in many applications such as copyright protection [2], content identification and multi-media database retrieval [3], object recognition [4], and remote sensing [5], where robustness, accuracy and computational efficiency are of paramount importance. Generally, image identification consists of two major steps, i.e. feature extraction and classification, in which an effective feature extraction algorithm is crucial for the entire recognition process, since a set of well-constructed features will make it much easier for correct recognition. This paper focuses on the development of an efficient and robust feature extraction algorithm. In the following, we will discuss the main challenges in developing such algorithms.

B. Main Challenges

Copyright protection of digital images is challenging as pictures of the same image will look very different taken from different angles, distances and lighting conditions. More importantly, it is often necessary to scan through a huge number of images on the Internet to detect a possible bootlegged image. For instance, more than 4 billion photos are available on Photobucket as of October 2011. In addition, Facebook alone publishes around 250 million photos per day [6] hosting 90 billion images in total as of January 2011. Flickr also hosts over 5 billion uploaded images as of September 2010 [7]. It is therefore crucial to develop a computationally efficient algorithm for invariant feature extraction with a very low false positive rate.

This paper aims to develop a robust yet computationally efficient feature extraction algorithm for copyright protection. The main idea is to find an optimal combination of the functionals as well as the optimal number of projections in Trace transform to achieve fast and robust feature extraction. The remainder of the paper is organized as follows. Section II gives a brief overview of related work. The Trace transform is depicted in Section III. Section IV introduces the concept of Pareto-optimal solutions and describes the proposed evolutionary algorithm. Experimental results for performance evaluations are presented in Section V. Section VI summarizes the paper and discusses the future work.

II. RELATED WORK

Feature extraction means accessing the relevant distinguishing information directly from the image [8]. For robust image recognition, extracted features should be insensitive to variations in geometric transformations such as rotation, scale and translation (RST). Features derived from different samples of the same image class should be similar. Meanwhile, features derived from samples of different image classes should considerably differ from each other [9]. Extracting invariant features from the same image class is a major challenge to image recognition. In the following, we discuss some feature extraction techniques related to the present work.

A. Moment Invariants

Over the past decades, image moments were widely investigated to construct invariant features. A number of different terms used for image moments are broadly defined describing pixel distribution in an image. Image moment is a tool for providing a solution for the geometric invariant problem [10]. Early work on image moments was undertaken by Hu [11] in 1962. In that work, the author first introduced a fundamental theorem of regular moment invariants for pattern recognition, also called geometric moments.

The 2-D moment of order $(p + q)$ of a digital image $f(x, y)$ of size $M \times N$ is defined as [12]

$$m_{pq} = \sum_{x=1}^M \sum_{y=1}^N x^p y^q f(x, y) \quad (1)$$

where m_{pq} is the moment of order $(p+q)$ $\{p+q = 0, 1, 2, \dots\}$.

The corresponding central moment of order $(p+q)$ is defined as

$$\mu_{pq} = \sum_{x=1}^M \sum_{y=1}^N (x - \bar{x})^p (y - \bar{y})^q f(x, y), \quad (2)$$

where $\bar{x} = \frac{m_{10}}{m_{00}}$ and $\bar{y} = \frac{m_{01}}{m_{00}}$ are the center of the gravity coordinates. Central moments in (2) are invariant to translation.

Hu [11] derived seven moments of order three independent of rotation, scale and translation of objects, which were revised later by Reiss [13]. Further, Flusser and Suk [14] derived four affine moment invariants as a generalization of previous work [15], which are widely used in literature.

Other invariant feature extraction techniques manipulate the transform domain, such as Fourier transform, wavelet transform, Radon transform and Trace transforms. In the following, we will focus on the last two transforms for their elegant properties they offer for invariant feature extraction.

B. Radon Transform

Radon transform has been widely used in digital image analysis. Johann Radon first introduced the transform in 1917 [16]. It is a useful tool in computed tomography medical imaging (CT scanner) to capture the directional feature of an image and enables the implementation of effective detection algorithm robust to noise [17][18].

Radon transform is determined by line integrals of an image $f(x, y)$ defined on the xy plane A and projected along straight lines of length ρ in a different angle θ . The two parameters ρ, θ characterize each line to represent a new 2D image in Radon space $R\{f(x, y)\}$ denoted as

$$R\{f(x, y)\} = \iint_A f(x, y) \delta(\rho - x \cos \theta - y \sin \theta) dx dy \quad (3)$$

$$\text{where } \delta(u) = \begin{cases} 1 & \text{if } u = 0 \\ 0 & \text{otherwise} \end{cases}$$

Over the last two decades, the Radon transform has received increased attention for image analysis and pattern recognition. It was shown [18] that Radon transform of the noise is constant and equal to the mean value of the noise. Therefore, it is insensitive to zero-mean white noise in the image. Due to its invertible property and robustness to noise, Radon transform was combined with other methods and transforms such as moment invariants, Fourier and wavelet transforms. A considerable amount of literature has been published on Radon transform. It has been successfully applied to a variety of image processing tasks such as texture analysis [19], robust watermarking [20] and iris image identification [21].

Inspired from Radon transform, a new algorithm as a tool for invariant feature extraction called Trace transform was proposed by Kadyrov and Petrou [22][23]. As our work is based on Trace transform, we will describe this algorithm separately in Section III.

III. THE TRACE TRANSFORM

In this section, we discuss briefly Trace transform and the theory of triple feature extraction. Trace transform was first proposed by Kadyrov and Petrou [22][23] where they presented the theory of invariant triple feature extraction. It has been applied to image registration [24] by defining the properties of the algorithm to identify the rotation and translation parameters between two images. Trace transform has received much interest in the field of invariant image analysis. It has been used in image analysis such as image database retrieval [22], texture classification [25], face authentication [26], and characters recognition [27].

Trace transform is an alternative representation of an image. It is obtained by tracing the image with straight lines and calculating a functional called "Trace" T over

parameter t on the pixel value along the straight lines. Each line is characterized by a length ρ and an angle θ , see Fig. 1(a), to form another 2D image with new coordinates θ and ρ , refer to Fig. 1(b) for an example. Each column of the trace matrix represents the tracing values for all lines oriented at that angle. Different transforms can be obtained by using different Trace functionals. Fig. 1(b) depicts an image and its Trace transform produced by using the first Trace functional in Table I (the integral of a function). In this case, the Trace transform is identical to Radon transform.

By applying a second functional called “Diametric” D along columns of Trace transform (along parameter ρ), a sequence of numbers is created. Finally, a third functional called “Circus” C is applied on the final string of numbers (over parameter θ) to produce a scalar value “real number”. This number termed a triple feature denoted by Π it can be a unique identifier for the image [28]. Fig. 2 summarizes the main steps for producing the triple feature. For robust image identification, the triple feature of an image should be very close to the triple feature of the distorted version of the same image. On the other hand, the triple feature of two different images should differ as much as possible. The plot of the diametric function produced by applying functional D to the trace matrix is depicted in Fig. 3. A rotation of the image is equivalent to a shift in the diametric function over parameter θ . One can easily observe that the curves are retained for the same image. However, there is a clear difference between the curves of the different images. The dotted curve in Fig. 3 represents a shift (demonstrated by an arrow) to the diametric plot of the original image (solid line).

TABLE I. LIST OF SOME TRACE FUNCTIONALS

No.	Functional	Description
1	$\int f(t) dt$	Radon Transform
2	$\left(\int f(t) ^p dt \right)^q$	p-Norm, $p = 0.5, q = 1/p$
3	$\int f(t)' dt$	Gradient
4	$\max(f(x))$	Maximum of absolute of the function

Figure 1. The Trace Transform. (a) The Trace parameters [22]. (b) An image and its Trace Transform.

Figure 2. A flowchart showing the main steps to construct the triple feature using the Trace algorithm.

Figure 3. Two different images and their corresponding diametric functions. The solid lines represent diametric of the original images and dotted lines represent a rotation in the original images by 90°.

One interesting property of Trace transform is that it has various invariant functionals that can be applied to produce different transforms [22][27]. Therefore, it can produce unlimited number of features. Consequently, depending on which functionals are used, different features can be obtained to describe the image invariant to general distortion such as rotation, scale and translation. Those features do not necessarily have a meaning in terms of human visual perception but are mathematically interpretable. Furthermore, various invariants can also be determined depending on the appropriate circus functional. Hence, the algorithm depends heavily on the functional combination T , D and C . In practice, these functionals are selected heuristically to achieve certain invariance. However, there is no theory for the selection of the correct combination of the Trace functionals T , D , and C for a given number of projections θ so that a good discrimination is achieved to identify different images. Consequently, selecting the best combination of these three functionals to produce a robust identifier is a challenging task. In this work, we adopted an evolutionary algorithm to tackle this problem.

In the next section, we will introduce a multi-objective evolutionary algorithm to produce a set of optimum triple features that describe the image.

IV. EVOLUTIONARY TRACE TRANSFORM (ETT)

Evolutionary Algorithms (EA) are powerful tools for finding the optimum solutions to complex problems. Although much research has been carried out on the Trace transform, little work has been done that applies machine learning and evolutionary computation to Trace transform. However, few exceptions exist that apply the machine learning techniques to fine tune Trace transform. In [29] a reinforcement learning algorithm was applied to the weighted Trace transform (WTT) to find the optimal threshold in the WTT space to minimize the within-class variance. In addition, Liu and Wang [30] have used more than one Trace functional and the Principle Component Analysis (PCA) to extract the hybrid trace features. The authors applied a genetic algorithm to calculate a single objective scalar that corresponds to each Trace functional for face recognition. The methods perform well, to some extent, in image retrieval and/or face recognition.

However, the methods above are computationally expensive since multiple Trace functionals need to be calculated for each image. Furthermore, the number of projections to trace the image is kept constant to perform the transform at all 360 directions. Heuristically, finding the optimal projections with a combination of the three functionals is a good idea to produce a balance between recognition accuracy and computational speed.

In most optimization methods, only a scalar cost function is optimized. However, it has been shown [31] that more than one objective should be considered in most optimization and learning problems. For example, feature extraction can be considered as a bi-objective optimization

problem, where, both within-class variance and the between-class variance should be taken into account.

The within-class variance S_w and between-class variance S_b are defined as follows [32]:

$$\begin{aligned} S_w &= \sum_{k=1}^K \sum_{j=1}^{N_k} (x_j^k - \mu_k)^2 \\ S_b &= \sum_{k=1}^K (\mu_k - \mu)^2 \end{aligned} \quad (4)$$

where $\mu_k = \frac{1}{N_k} \sum_{j=1}^{N_k} x_j^k$, $\mu = \frac{1}{K} \sum_{k=1}^K \mu_k$ and

K : Number of classes,

N_k : Number of samples in class k ,

μ_k : Mean of class k ,

x_j^k : The j^{th} sample of class k ,

μ : Mean of all classes.

In the following, the main components of the evolutionary algorithm will be presented.

1) *Chromosome*: Each chromosome in the evolutionary algorithm encodes four integer parameters, namely, the three functionals: trace, diametric and circus, and the number of projections.

2) *Population*: The initial population is generated randomly taking the constraints on the design variables into account.

3) *Fitness*: The evolutionary algorithm is set to minimize the following two objectives:

$$\begin{aligned} f_1 &= S_w \\ f_2 &= 1/(S_b + \epsilon) \end{aligned} \quad (5)$$

where S_w and S_b are the within-class variance and between class variance defined in (4), ϵ is a small quantity to avoid division by zero.

4) *Selection*: The selection strategy is based on the elitist non-dominated sorting method [33][34], which is composed of four main steps. First, the parent and offspring populations are combined. Second, all individuals in the combined population are assigned a Pareto front number and a crowding distance. In Pareto front assignment, the non-dominated solutions in the combined population are assigned a rank 1, which belongs to the first non-dominated front. These individuals are removed temporally from the population, and the non-dominated individuals in the rest of the population are identified, which form the second non-dominated front of the population and are assigned a rank 2.

This procedure repeats until all individuals in the combined population are assigned with a rank from 1 to r , assuming that r non-dominated fronts can be identified in total. Third, all individuals are then sorted according to the assigned Pareto front number in an ascending order and individuals having the same Pareto front number are sorted according to the crowding distance in a descending order. Finally, the top N_p individuals, where N_p is the population size are selected and passed to the next generation. The reader is referred to [33] for details about non-dominated sorting and crowding distance calculation.

5) *Crossover and mutation*: In generating offspring, two solutions are chosen at random from the mating pool to exchange a portion of the string to produce new solutions. A uniform crossover is adopted with crossover probability P_c . Mutation is applied to the offspring after crossover. It is not necessary that good solutions will be created through crossover and mutation. However, only better solution will survive through the selection operator.

At the end of the evolution, the final non-dominated solutions are analyzed and used as feature extraction on unseen images.

V. EXPERIMENTAL RESULTS

In the experiment, we use the multi-objective evolutionary algorithm described in Section IV to search for the best combination of the Trace functionals and the optimal number of projections. We use 14 trace functionals (T), six diametric functionals (D) and six circus functionals (C). Each functional combination will be applied with a projection angles between 180 and 360 degrees. A smaller projection angle for a given functionals means lower computational complexity and therefore a faster recognition speed. The parameters used in the experiments are listed in Table II.

A set of trace, diametric and circus functionals can be found in [22][28][35] and we list four functionals in Table I in Section III. The Trace algorithm is first run once offline to calculate the tracing parameters for each line in the image. Then the main program reads in the stored parameters to produce the transform. This will help reduce the run-time of the algorithm. The reader is referred to [36] for a tutorial on implementation of the Trace Algorithm.

The multi-objective evolutionary algorithm is run to search for Pareto-optimal solutions for five different fish images, each having three distorted versions (rotated, scaled and translated) to form 20 images in total (see Fig. 10). Then, we pick out a few solutions from the Pareto-optimal set and test them on unseen images from the database. Fig. 4 shows three non-dominated fronts near the Pareto-optimal front. The Pareto-optimal front in the final generation is depicted in Fig. 5. It is found that all Pareto-optimal solutions shown in Fig. 5 can produce satisfying results. In the following experiments, we pick two solutions having a projection number smaller than 360 for less complexity (indicated by the arrows in Fig. 5) and compare their

performance with that of the Trace algorithms reported in [22].

It can be seen from Fig. 6 that diametric plots reveal the invariance in image features. The scaled and translated versions of the image almost exactly match the original image and the rotated version of the image represented in a shift in θ . However, it maintains the shape of the diametric plot of the original image.

The non-overlapped features are shown in Fig. 7. The figure shows the scatter plot of five image classes with one-dimensional feature only. In our work, one-dimensional feature means that we run the algorithm only once as illustrated earlier in Fig. 2. Although one-dimensional features show excellent variance and class scatter, it is worth drawing 2D features to compare them to the features reported in [22]. Here, it is important to normalize our features for comparison purpose. In Fig. 8, we re-plot the scatters of the five image classes in [22]. In Fig. 9, we plot the scatter of the same images using our optimal features. Comparing the results in Figs. 8 and 9, one can easily observe that features obtained in our work exhibit more stability than those in [22]. We should also mention that features in Fig. 8 are constructed by taking ratios of two different triple features, i.e. four triple features in total are required to draw the 2D features [22]. However, our results are obtained using only three functionals, forming a triple feature i.e. only two triple features in total to draw the 2D features in Fig. 9. Thus, our method is faster and more robust compared to the results presented in [22], [23] and [37].

The experiments were performed on Intel®Core™2Duo 3.1GHz processor with 3GB RAM using Visual C++ compiler. The time for optimization using the NSGA-II took around eight hours for 100 generations and the Trace algorithm itself needs only a few seconds due to our efficient implementation of the Trace algorithm. Table III depicts the value of two triple features Π_1 and Π_2 .

To evaluate the effectiveness of the proposed ETT, the ratio of the scatters obtained from this algorithm is compared to that obtained by the existing variants of the Trace transform. Table IV summarizes the value of S_w/S_b obtained by our method and the method used in [22] for each feature Π_1 to Π_5 . As we mentioned earlier, features Π_1 to Π_5 (present in Tables IV, and their pairs in Tables V and VI) are different in our method from those in [22]. As we can see, the ratio of S_w/S_b resulting from ETT is considerably better than the ratio obtained from features presented in [22], where the ratio is lower in our method than the minimum obtained by any feature in [22].

It is also worth calculating the distance between ratios of the class scatter obtained by each pairs of features by

$$\sigma = \sqrt{\left(\frac{S_w}{S_b}\right)_{\Pi_i}^2 + \left(\frac{S_w}{S_b}\right)_{\Pi_j}^2} \quad (6)$$

where $i, j = 1, 2, 3, \dots, i \neq j$.

Tables IV, V and VI clearly demonstrate the advantage of the proposed algorithm over the existing ones.

We have also tested the performance of the optimized Trace algorithm on another five images and their modified versions, forming altogether 40 images of ten classes. Fig. 11 shows some of the extra images. We can notice from Fig. 12 that features of the ten images studied in our experiment have a very small within-class variance and a large between-class variance.

Figure 4. Three non-dominated fronts near the optimal front

TABLE II. PARAMETERS SETUP

Parameter	Value
Population size N_p	100
Mutation probability P_m	0.025
Crossover probability P_c	0.9
Number of generations	100
$\mathcal{E}$	10^{-5}

TABLE III. TRIPPLE FEATURES FOR FISH IMAGES

Fish	Triple Feature $\Pi_1 \times 10^3$			
	Version1	Version2	Version3	Version4
A	1.1954	1.1954	1.1992	1.1960
B	1.3419	1.3419	1.3421	1.3408
C	0.3186	0.3185	0.3184	0.3191
D	1.0041	1.0041	1.0044	1.0043
E	0.6039	0.6039	0.6038	0.6046
Fish	Triple Feature $\Pi_2 \times 10^3$			
	Version1	Version2	Version3	Version4
A	17.7069	17.7068	17.8490	17.7247
B	28.242	28.2416	28.363	28.2328
C	8.2693	8.26772	8.27564	8.26987
D	17.4054	17.4057	17.4772	17.4044
E	12.4433	12.4447	12.5052	12.4508

Figure 5. Non-dominated solutions as a Pareto-optimal front in the objective space.

Figure 6. Diametric value of two different images and their modified versions. Notice the exact match among all shapes and the shift in the rotated version of the images in both (a) and (b).

Figure 7. Plotting of 1D invariant features for the images in Fig. 10 extracted using our method.

Figure 8. Re-plotting of best 2D invariant features reported in [22] for the images in Fig. 10 (5×4 images).

Figure 9. Plotting of normalized 2D invariant features for the images in Fig. 10 (5x4 images).

Figure 10. Fish database. Rows (A-E) represent four different versions of the same image [22].

Figure 11. More images from the database.

TABLE IV. COMPARISON OF S_w/S_b BETWEEN FEATURES IN [22] AND ETT

	S_w/S_b	
	Features in [22] ($\times 10^{-3}$)	ETT ($\times 10^{-3}$)
Π_1	725.5	0.01677
Π_2	274.8	0.1439
Π_3	5.937	0.3387
Π_4	51.362	0.0437
Π_5	13.99	0.0066

TABLE V. THE DISTANCE σ FROM FEATURES IN [22]

2D Feature Combinations in [22] ($\times 10^{-3}$)			
(Π_1, Π_2)	(Π_5, Π_3)	(Π_5, Π_2)	(Π_4, Π_3)
284.2	15.2	275.2	51.7

TABLE VI. THE DISTANCE σ FROM FEATURES OBTAINED BY ETT

2D Feature Combinations obtained by ETT ($\times 10^{-3}$)			
(Π_1, Π_2)	(Π_1, Π_3)	(Π_2, Π_4)	(Π_4, Π_5)
0.15	0.34	0.14	0.0439

Figure 12. Plotting of 2D invariant features from ETT for the images in Fig. 10 and Fig. 11 (10x4 images).

VI. CONCLUSION

We have adopted an evolutionary multi-objective optimization algorithm, NSGA-II for choosing an optimal combination of the functions as well as an optimal number of projections of the Trace algorithm, which is termed ETT. Our empirical results indicate that proposed ETT can produce very consistent within-class features and very different between-class features, which is of essential importance for robust and accurate image recognition. The use of pre-stored Trace parameters reduced the run-time of

the Trace algorithm considerably. Empirical results show that the optimized Trace algorithm is able to produce highly invariant features for different versions of the same image. Meanwhile, features from different images are sufficiently different even with one-dimensional features. Future work will be to verify the performance of ETT on more images from different image databases.

ACKNOWLEDGMENT

Thanks to Prof. M. Petrou for providing the fish images database.

REFERENCES

- [1] I. Cox, M. Miller, J. Bloom, J. Fridrich and T. Kalker, *Digital Watermarking and Steganography*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc, 2007.
- [2] K. Naoe and Y. Takefuji, "Damageless image hashing using neural network," in *International Conference of Soft Computing and Pattern Recognition (SoCPaR)*, 2010, 2010, pp. 442-447.
- [3] A. Wong and W. Bishop, "Robust invariant descriptor for symbol-based image recognition and retrieval," in *International Conference on Semantic Computing, 2007. ICSC 2007*, 2007, pp. 637-644.
- [4] S. Lee, J. Oh, J. Park, J. Kwon, M. Kim and H. Yoo, "A 345 mW Heterogeneous Many-Core Processor With an Intelligent Inference Engine for Robust Object Recognition," *IEEE Journal of Solid-State Circuits*, vol. 46, pp. 42-51, 2011.
- [5] G. J. Scott, M. N. Klaric, C. H. Davis and C. Shyu, "Entropy-Balanced Bitmap Tree for Shape-Based Object Retrieval From Large-Scale Satellite Imagery Databases," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, pp. 1603-1616, 2011.
- [6] "Facebook Statistics," <http://www.facebook.com/press/info.php?statistics>. 2011.
- [7] "5 billionth photo uploaded to Flickr - CNN ". Date accessed: December 2011, http://articles.cnn.com/2010-09-20/tech/flickr.5.billion_1_photo-sharing-site-flickr-facebook?_s=PM:TECH.
- [8] J. S. Seo, J. Haitsma, T. Kalker and C. D. Yoo, "A robust image fingerprinting system using the Radon transform," *Signal Process Image Commun*, vol. 19, pp. 325-339, 4, 2004.
- [9] G. Y. Chen and B. Kegl, "Feature extraction using radon, wavelet and fourier transform," in *IEEE International Conference on Systems, Man and Cybernetics, 2007. ISIC. 2007*, pp. 1020-1025.
- [10] M. Alghoniemy and A. H. Tewfik, "Geometric Invariance in image watermarking," *IEEE Transactions on Image Processing*, vol. 13, pp. 145-153, 2004.
- [11] M. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, vol. 8, pp. 179-187, 1962.
- [12] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. New Jersey: Pearson Prentice Hall, 2008.
- [13] T. H. Reiss, "The revised fundamental theorem of moment invariants," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 830-834, 1991.
- [14] J. Flusser and T. Suk, "Pattern recognition by affine moment invariants," *Pattern Recognit*, vol. 26, pp. 167-174, 1, 1993.
- [15] S. Maitra, "Moment invariants," *Proc IEEE*, vol. 67, pp. 697-699, 1979.
- [16] S. R. Deans, *The Radon Transform and some of its Applications*. Krieger Publishing Company, 1983.
- [17] K. Jafari-Khouzani and H. Soltanian-Zadeh, "Rotation-invariant multiresolution texture analysis using Radon and wavelet transforms," *IEEE Transactions on Image Processing*, vol. 14, pp. 783-795, 2005.
- [18] D. V. Jadhao and R. S. Holambe, "Feature extraction and dimensionality reduction using radon and fourier transforms with application to face recognition," in *International Conference on Computational Intelligence and Multimedia Applications, 2007. 2007*, pp. 254-260.
- [19] G. Liu, Z. Lin and Y. Yu, "Radon Representation-Based Feature Descriptor for Texture Classification," *IEEE Transactions on Image Processing*, vol. 18, pp. 921-928, 2009.
- [20] H. Zhu, M. Liu and Y. Li, "The RST invariant digital image watermarking using Radon transforms and complex moments," *Digital Signal Processing*, vol. 20, pp. 1612-1628, 12, 2010.
- [21] Y. Zhou and A. Kumar, "Personal identification from iris images using localized radon transform," in *20th International Conference on Pattern Recognition (ICPR), 2010*, 2010, pp. 2840-2843.
- [22] A. Kadyrov and M. Petrou, "The Trace transform and its applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 811-828, 2001.
- [23] A. Kadyrov and M. Petrou, "The trace transform as a tool to invariant feature construction," in *Fourteenth International Conference on Pattern Recognition, 1998. Proceedings. 1998*, pp. 1037-1039 o.2.
- [24] A. Kadyrov and M. Petrou, "Linear transformation parameter estimation for fault detection," in *Fourteenth International Conference on Pattern Recognition, 1998. Proceedings. 1998*, pp. 550-552 o.1.
- [25] A. Kadyrov, A. Talepbour, M. Petrou, "Texture classification with thousands of features," in *13th British Machine Vision Conference*, Cardiff, UK, 2002, pp. 656-665.
- [26] S. Srisuk, M. Petrou, W. Kurutach and A. Kadyrov, "Face authentication using the trace transform," in *Proceedings. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. 2003*, pp. -305-312 o.1.
- [27] M. F. Nasrudin, M. Petrou and L. Kotoulas, "Jawi character recognition using the trace transform," in *Seventh International Conference on Computer Graphics, Imaging and Visualization (CGIV), 2010*, 2010, pp. 151-156.
- [28] M. Petrou and A. Kadyrov, "Affine invariant features from the trace transform," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 30-44, 2004.
- [29] S. Srisuk, R. Fooprateepsiri, M. Petrou, S. Waraklang and K. Sunat, "A general framework for image retrieval using reinforcement learning," in *The Image and Vision Computing 2003*, Massey University, New Zealand, 2003, pp. 36-41.
- [30] N. Liu and H. Wang, "Extraction of hybrid trace features with evolutionary computation for face recognition," in *IEEE Congress on Evolutionary Computation, 2007. CEC 2007. 2007*, pp. 2493-2500.
- [31] Y. Jin and B. Sendhoff, "Pareto-Based Multiobjective Machine Learning: An Overview and Case Studies," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 38, pp. 397-415, 2008.
- [32] A. M. Martinez and A. C. Kak, "PCA versus LDA," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 228-233, 2001.
- [33] K. Deb, *Multi-Objective Optimization using Evolutionary Algorithms*. England: John Wiley & Sons. Ltd, 2002.
- [34] K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, pp. 182-197, 2002.
- [35] M. F. Nasrudin, K. Omar and Zakaria, Choong-Yeun Liong and Mohamad Shanudin, "Invariant features from the trace transform for jawi character recognition," in *Proceedings of the 10th International Work-Conference on Artificial Neural Networks: Part II*, Salamanca, Spain, 2009, pp. 256-263.
- [36] M. Petrou and F. Wang, *A Tutorial on the Practical Implementation of the Trace Transform*, ser. Handbook of Texture Analysis. Imperial College Press, 2008, ch. 11, pp. 313-346.
- [37] M. Petrou and A. Kadyrov, "Features invariant to affine distortions from the trace transform," in *Proceedings. International Conference on Image Processing, 2001. 2001*, pp. 852-855 o.3.