2016 UKSim-AMSS 18th International Conference on Computer Modelling and Simulation

Pose Invariant Thermal Face Recognition Using AMI Moments

Naser Zaeri
Faculty of Computer Studies
Arab Open University
P.O. Box 3322 Safat 13033, Kuwait
e-mail: n.zaeri@aou.edu.kw

***Abstract*—Imaging in the visible spectrum demonstrates difficulties in recognizing the faces in conditions of varying illumination, especially under total darkness conditions. Further, the pose variations in such images add extra burden and heavy challenge on successful performance. As such, thermal face recognition has laid itself as a successful alternative solution and eventually has become an area of growing interest. In this paper, we present a new technique for thermal face recognition based on affine moment invariants (AMI) technique. AMI technique has become one of the most important shape descriptors. The technique will be implemented at the component level by dividing the face image into non-overlapped components. We anticipate that this approach will offer robustness against variability due to changes in localized regions of the faces. The new method will be tested on a new database comprising of images of different expressions with various severe poses, and were taken within different time-lapse. The experimental results have shown that the proposed technique offers high discriminability and performs efficiently, with Rank-1 successful rate of ~95% over the different poses.**

Keywords- face recognition; thermal image; moment invariants; pose variations

I. INTRODUCTION

Researchers found that it is important to attempt to understand the strategies that the biological system employs, as a first step towards eventually translating these strategies into machine-based algorithms. These observations provide useful hints that can be valuable to computer vision systems [1]. In our electronically inter-connected society, reliable and user-friendly face recognition and verification system is essential in many sectors of our life.

Currently, most researches on face recognition focus on visual images. Although considerable progress has been made in the domain of face recognition over the last decade, especially with the development of powerful methods, face recognition has shown to be not accurate enough in uncontrolled environments. Face recognition performances of a system can be degraded by many factors, including facial expression, head pose variation, occlusion and most importantly illumination changes [2].

Over the last few years, thermal IR imaging based face recognition has emerged as a promising complement to conventional visible spectrum based approaches [3]. Different objects emit different range of infrared energy according to their temperature and characteristics. The range of human face and body temperature is nearly the same and quite uniform. This provides a consistent thermal signature. IR cameras provide a measure of thermal emissivity from the facial surface, and their images are relatively stable under illumination variation. The anatomical information which is imaged by infrared technology involves subsurface features.

In this paper, we present a new technique for face recognition that exploits the statistical characteristics of a thermal image by the virtue of moment invariants. The statistical features of the images find a combination of multiple statistical patterns to produce a result that is enhanced in terms of information content for pattern recognition and classification. Moment invariants offer robustness against variability due to the changes in regions of the objects. The evaluation uses a database of thermal face images that has been developed in the Artificial Intelligence laboratory at the Arab Open University (AIAOU Database) [4]. The organization of the paper is as follows. A brief literature review is given in Section 2. A mathematical background of the proposed method is furnished in Section 3. The experimental results are discussed in Section 4. Finally, the paper is brought to a conclusion in Section 5.

II. LITERATURE REVIEW

Due to its physiology, a human face consists of "hot" parts that correspond to tissue areas that are rich in vasculature and "cold" parts that correspond to tissue areas with sparse vasculature [5]. Every living and non-living object at a finite temperature emits radiation, which can be captured by infrared cameras. Early studies by Socolinsky et al. in [6, 7] suggest that long-wave infrared imagery of human faces is not only a valid biometric, but superior to using comparable visible-light imagery. Prokoski et al. [8] anticipated the possibility of extracting the vascular network from thermal facial images and using it as a feature space for face recognition. However, they did not present an algorithmic approach for achieving this.

Bhowmik et al. [9] introduced the role of different IR spectrums and their applications. In their experimental work, they fused both thermal and visible images to enhance the recognition rate, as it is expected that the fusion process improves the overall performance of the system. They tested their method on IRIS and Terravic databases. The images of both databases were taken in one session. Guzman et al. [10] discussed a thermal imaging framework that consolidates the

978-1-5090-0888-9/16 $31.00 © 2016 IEEE
DOI 10.1109/UKSim.2016.42

steps of feature extraction through the use of morphological operators and registration using the linear image registration tool. The matching showed an average accuracy of 88.46% for skeletonized signatures and 90.39% for anisotropically diffused signatures.

Papakostas et al. [11] employed orthogonal moment features as pattern descriptors of thermal face images. In this work, they have utilized granular kNN lattice computing techniques. The reported experimental results are promising. Unfortunately, they tested their method on a subset of Terravic Facial IR database consisting of 700 images only. Seal et al. [12] focused on preprocessing the thermal face image before feature extraction. They have noticed that some holes are created in the face area due to uneven distribution of thermal information which is nothing but variation in temperature statistics. These temperature statistics have been excluded in the binarization process. They implemented Gappy-PCA approach to store the statistical information of missing temperature zones.

Li et al. [13, 14] used features based on local binary patterns (LBP) extracted from infrared images. They applied their algorithm in an active setting which uses strong NIR light-emitting diodes which is coaxial with the direction of the camera. However, it is unsuitable for uncooperative user applications or outdoor use due to the strong NIR component of sunlight.

Abas and Ono [15] proposed the implementation of moment invariants; with respect to centroid point obtained from frontal mugshot images. The system decomposes a background filtered thermal image into 4 thermal regions via 3-valued threshold method. In their work, they have only used the first moment invariant, I1, from Hu's classical moment invariants. However, they have not explained the virtue behind such decomposition. Also, the database used is a small one and the authors have used only the frontal images with slight deviation in pose. No serious or challenging parameters were considered in their evaluation. The best recorded result was 92.5%. A different attempt at extracting invariant features which exploits the temperature differential between vascular and non-vascular tissues was proposed by Wu et al. [16] and Xie et al. [17].

III. AFFINE MOMENT INVARIANTS

The approach using invariant features is based on describing the objects by a set of measurable quantities called invariants that are insensitive to particular deformations and that provide enough discrimination power to distinguish objects belonging to different classes [18]. From a mathematical point of view, invariant I is a functional defined on the space of all admissible image functions that does not change its value under degradation operator D, i.e. that satisfies the condition $I(f) = I(D(f))$ for any image function f. Usually, one invariant does not provide enough discrimination power and several invariants $I_1, \ldots, I_n$ must be used simultaneously. This will lead to having an invariant vector. In this way, each object is represented by a n-dimensional vector space called *feature space* or *invariant space*. Moments are scalar quantities used to characterize a function and to capture its significant features. From the mathematical point of view, moments are "projections" of a function onto a polynomial basis. Depending on the polynomial basis used, various systems of moments can be recognized.

Geometric moment of order (p+q) for a two dimensional discrete function is computed using (1),

$$m_{pq} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} x^p y^q f(x,y) \qquad (1)$$

where $f(x, y)$ is the image function and M, N are image dimensions. The image function can be exactly reconstructed from the set of its moments. Invariance to translation can be achieved simply by seemingly shifting the object such that its centroid coincides with the origin of the coordinate system or, vice versa, by shifting the polynomial basis into the object centroid. In the case of geometric moments, we have the so-called central geometric moments

$$\mu_{pq} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} (x-\bar{x})^p (y-\bar{y})^q f(x,y) \qquad (2)$$

where $f(x, y)$ is the image function and M, N are image dimensions, and $\bar{x} = m_{10}/m_{00}$, $\bar{y} = m_{01}/m_{00}$ are the coordinates of the object centroid.

In a vast majority of practical cases, pictures of 3D scenes or planar scenes are taken place arbitrarily in a 3D environment. In such cases, 3D objects and structures are represented by their projections onto a 2D plane since photography is a 2D medium [19]. Because of this, we often face object deformations that are beyond the translation-rotation-scaling model. AMIs play a very important role in moment-based theories and can be considered as a crucial tool in these aspects.

Affine transformation is a general linear transform of spatial coordinates of the image, which can (under certain circumstances) approximate the 3D object representation. They are invariant with respect to affine transform of the spatial coordinates. The AMIs can be derived in several ways that differ from each other in the mathematical tools used. One way to derive these moments is the algebraic invariants. Further, graph theory, tensor algebra, partial differential equations, and derivation via image normalization are different techniques that researchers have proposed in this domain. Suk and Flusser have presented a complete and independent list of AMIs in [20]. The first nine independent moments are explicitly given in equations (3–11). The human head, with its face in different poses, can be considered as a 3D shape where the implementation of AMIs may lead to a suitable representation for face recognition. In this work, we propose a feature vector I consisting of the 9 different moments described in equations (3-11), as shown in equation 12. The Euclidean distance (L_2 norm) is used as the system classifier and is given by (13), for vectors **a** and **b** both of d dimensions.

$$I_1 = (\mu_{20}\mu_{02} - \mu_{11}^2)/\mu_{00}^4 \tag{3}$$

$$I_2 = (-\mu_{30}^2\mu_{03}^2 + 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} - 4\mu_{30}\mu_{12}^3 - 4\mu_{21}^3\mu_{03} + 3\mu_{21}^2\mu_{12}^2)/\mu_{00}^{10} \tag{4}$$

$$I_3 = (\mu_{20}\mu_{21}\mu_{03} - \mu_{20}\mu_{12}^2 - \mu_{11}\mu_{30}\mu_{03} + \mu_{11}\mu_{21}\mu_{12} + \mu_{02}\mu_{30}\mu_{12} - \mu_{02}\mu_{21}^2)/\mu_{00}^7 \tag{5}$$

$$\begin{aligned} I_4 = (&-\mu_{20}^3\mu_{03}^2 + 6\mu_{20}^2\mu_{11}\mu_{12}\mu_{03} - 3\mu_{20}^2\mu_{02}\mu_{12}^2 - 6\mu_{20}\mu_{11}^2\mu_{21}\mu_{03} - 6\mu_{20}\mu_{11}^2\mu_{12}^2 + 12\mu_{20}\mu_{11}\mu_{02}\mu_{21}\mu_{12} - 3\mu_{20}\mu_{02}^2\mu_{21}^2 \\ &+ 2\mu_{11}^3\mu_{30}\mu_{03} + 6\mu_{11}^3\mu_{21}\mu_{12} - 6\mu_{11}^2\mu_{02}\mu_{30}\mu_{12} - 6\mu_{11}^2\mu_{02}\mu_{21}^2 + 6\mu_{11}\mu_{02}^2\mu_{30}\mu_{21} - \mu_{02}^3\mu_{30}^2)/\mu_{00}^{11} \end{aligned} \tag{6}$$

$$I_5 = (\mu_{40}\mu_{04} - 4\mu_{31}\mu_{13} + 3\mu_{22}^2)/\mu_{00}^6 \tag{7}$$

$$I_6 = (\mu_{40}\mu_{22}\mu_{04} - \mu_{40}\mu_{13}^2 - \mu_{31}^2\mu_{04} + 2\mu_{31}\mu_{22}\mu_{13} - \mu_{22}^3)/\mu_{00}^9 \tag{8}$$

$$I_7 = (\mu_{20}^2\mu_{04} - 4\mu_{20}\mu_{11}\mu_{13} + 2\mu_{20}\mu_{02}\mu_{22} + 4\mu_{11}^2\mu_{22} - 4\mu_{11}\mu_{02}\mu_{31} + \mu_{02}^2\mu_{40})/\mu_{00}^7 \tag{9}$$

$$\begin{aligned} I_8 = (&\mu_{20}^2\mu_{22}\mu_{04} - \mu_{20}^2\mu_{13}^2 - 2\mu_{20}\mu_{11}\mu_{31}\mu_{04} + 2\mu_{20}\mu_{11}\mu_{22}\mu_{13} + \mu_{20}\mu_{02}\mu_{40}\mu_{04} - 2\mu_{20}\mu_{02}\mu_{31}\mu_{13} + \mu_{20}\mu_{02}\mu_{22}^2 + 4\mu_{11}^2\mu_{31}\mu_{13} \\ &- 4\mu_{11}^2\mu_{22}^2 - 2\mu_{11}\mu_{02}\mu_{40}\mu_{13} + 2\mu_{11}\mu_{02}\mu_{31}\mu_{22} + \mu_{02}^2\mu_{40}\mu_{22} - \mu_{02}^2\mu_{31}^2)/\mu_{00}^{10} \end{aligned} \tag{10}$$

$$\begin{aligned} I_9 = (&\mu_{20}\mu_{30}\mu_{12}\mu_{04} - \mu_{20}\mu_{30}\mu_{03}\mu_{13} - \mu_{20}\mu_{21}^2\mu_{04} + \mu_{20}\mu_{21}\mu_{12}\mu_{13} + \mu_{20}\mu_{21}\mu_{03}\mu_{22} - \mu_{20}\mu_{12}^2\mu_{22} - 2\mu_{11}\mu_{30}\mu_{12}\mu_{13} \\ &+ 2\mu_{11}\mu_{30}\mu_{03}\mu_{22} + 2\mu_{11}\mu_{21}^2\mu_{13} - 2\mu_{11}\mu_{21}\mu_{12}\mu_{22} - 2\mu_{11}\mu_{21}\mu_{03}\mu_{31} + 2\mu_{11}\mu_{12}^2\mu_{31} + \mu_{02}\mu_{30}\mu_{12}\mu_{22} \\ &- \mu_{02}\mu_{30}\mu_{03}\mu_{31} - \mu_{02}\mu_{21}^2\mu_{22} + \mu_{02}\mu_{21}\mu_{12}\mu_{31} + \mu_{02}\mu_{21}\mu_{03}\mu_{40} - \mu_{02}\mu_{12}^2\mu_{40})/\mu_{00}^{10} \end{aligned} \tag{11}$$

$$I = [m_{00}, m_{10}, m_{01}, \mu_{11}, \mu_{12}, \mu_{21}, \mu_{22}, \mu_{30}, \mu_{03}, \mu_{40}, \mu_{04}] \tag{12}$$

$$D(\mathbf{a}, \mathbf{b}) = \left(\sum_{k=1}^{d} (a_k - b_k)^2 \right)^{1/2} \tag{13}$$

IV. Experimental Work

We have built a database consisting of 1500 images for 20 different subjects taken in different sessions. We used the Infrared Camera ETIP 7320 which includes thermal infrared imaging radiometer using a micro-bolometer 320×240 focal plane array and a Vanadium Oxide technology base, ensuring high efficient thermal and spatial resolution. The images are of varying poses and expression conditions. Participants were asked to express three different emotions with their faces: neutral, anger, and smiling. For every expression, we have acquired five different images at five varying poses: 0°, 45°, 90°, 135°, and 180°. Image at 0° represents a person looking in the direction of his/her right-hand shoulder. Further, image at 90° represents the frontal case view, whereas an image at 180° represents a person looking in the direction of his/her left-hand shoulder. Further, acquisitions were held at different times across a number of different weeks. Also, the database consists of males and females from various ethnic backgrounds. Examples of face images from the database are shown in Fig. 1.

In our work, the proposed method rely on measuring certain characteristics (moments) from a certain set of images. We refer to this set as a training set which is divided into five subsets according to the five poses (0°, 45°, 90°, 135°, and 180°), where each pose consists of five images for every person (these images have different expression variations and have been *randomly* chosen). The rest 10 images (for a specific person) from each pose have been used as the testing set (Again; these 10 images contain variations in expression). The test image is first categorized to a certain pose (one of the five described different poses). Then, it is tested against the different available classes (20 classes) corresponding to that specific pose. So, the first step is to associate the probe image at hand to one of the five available pose subsets. Hence, we perform data clustering to associate the data subsets for individuals with similar thermal appearance in a particular range of views. This will allow the specialization of moments trained only on this particular cluster. The face pose is found by locating the nose and mouth from the thermal signature.

Unlike the visible spectrum case, in which background clutter is often significant and in which face segmentation can be a difficult task, face segmentation in thermal images is in most cases far simpler. We have created a provisional segmentation map by declaring all pixels with values within a range between two thresholds, T_{min} and T_{max}, as belonging to the face region (i.e. foreground) and all others as background. Eventually all images were cropped to size 45×40.

We have implemented the new proposed technique on the face images at the "component level". At the component-based approach, the face image is divided into components where the local characteristics and features are combined together using a certain fusion method. Each moment is trained on a determined cluster (component) of thermal images in the training database. Then, the local features are combined at a second stage to decide whether the input face image belongs to a given class. We propose dividing the image into 16 components. These components are of equal size and are non-overlapped. The results obtained from the similarity measures from the feature vectors for the different number of components are fused together to achieve the final similarity score. In our experiments, we have implemented "*voting*" fusion. Fig. 2 shows a thermal image for one subject divided into 16 components. Fig. 3 shows the recognition rates versus face pose for the proposed method. As we can see from the graph, the proposed approach has resulted in outstanding results with a successful recognition rate ranging from 93% to 95% over the different poses. These results show that the performance of the proposed method is stable and efficient even with images of large poses. Moreover, the challenge related to the existence of images with eyeglasses has been very well coped with.

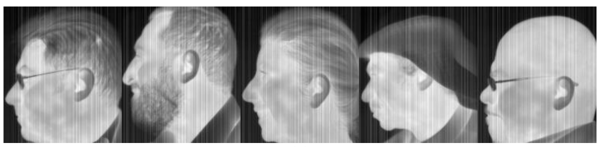
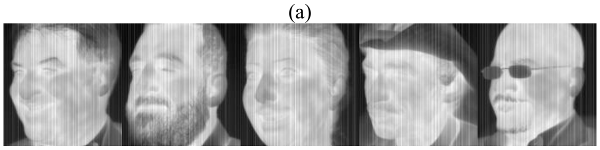
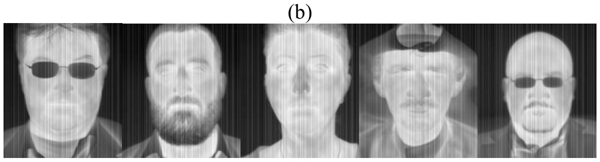
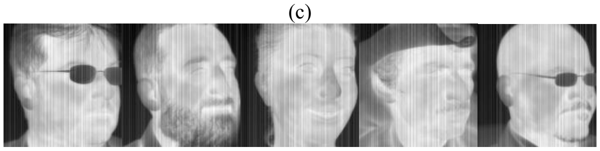
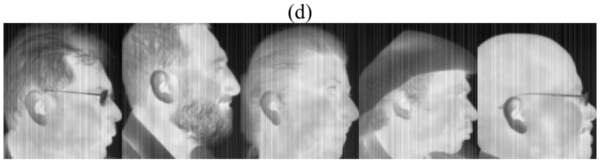

(a)

(b)

(c)

(d)

(e)

Figure 1. Examples of thermal images from the AIAOU database with different poses and expressions: a) 0°, b) 45°, c) 90°, d) 135°, and e) 180°.

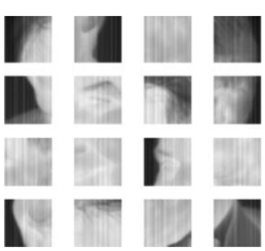


Figure 2. An example of the 16-component image for one subject with pose = 45°.

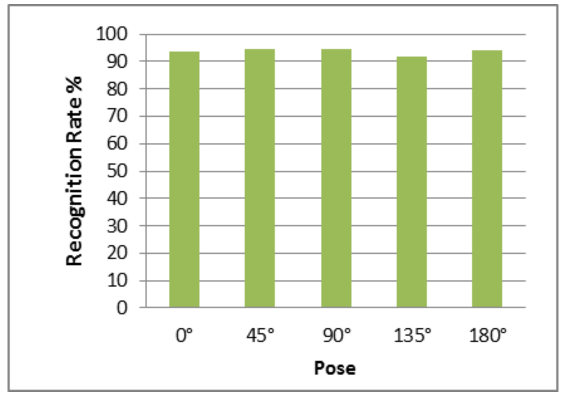


Figure 3. Recognition rate vs. face pose using 16-component approach.

V. CONCLUSION

In this paper, we have presented a new technique for face recognition based on statistical calculations of thermal images. We proposed the use of AMI moments which become one of the most important shape descriptors. The proposed technique has been implemented at the component level by dividing the face image into 16 components. The component-based AMI moments technique has utilized the local representations efficiently. These representations offered robustness against variability due to the changes in poses of a face image. The features at the local level are easier for estimating the rotations and misalignment. As the experimental results reveal, the technique is stable and can deal with the pose variations efficiently with recognition rates of ~95% over the different poses.

Finally, it should be noted that the component-based approach should consider various parameters in order to be successful. One of these parameters is the suitable number of components.

ACKNOWLEDGMENT

The author would like to express his gratitude and grateful appreciation to the Kuwait Foundation for the Advancement of Sciences (KFAS) for fully funding this project under the code: 2013-1510-01.

REFERENCES

[1] Huang, Z. Lei, M. Fan, X. Wang, "Regularized discriminative spectral regression method for heterogeneous face matching," IEEE Transactions on Image Processing, Vol. 22, No. 1, pp. 353-362, 2013.

[2] F. Nicolo, NA. Schmid, "Long range cross-spectral face recognition: Matching SWIR against visible light images," IEEE Transactions on Information Forensics and Security, Vol. 7, No. 6, pp. 1717-1726, 2012.

[3] B.A. Rajoub, R. Zwiggelaar, Thermal Facial Analysis for Deception Detection, IEEE Transactions on Information Forensics and Security, Volume: 9, Issue: 6, pp. 1015 – 1023, 2014.

[4] N. Zaeri, "Component-based Thermal Face Recognition," British Journal of Applied Science & Technology, vol. 4, no. 6, pp. 945-966, 2014.

[5] C. San Martin, , P. Meza, S. Torres, and R. Carrillo, "Improved infrared face identification performance using nonuniformity

correction techniques," Lecture Notes on Computer Science, vol. 5259, pp. 1115-1123, 2008.

[6] D. Socolinsky, and A. Selinger, "Thermal face recognition in an operational scenario," Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2004, vol. 4, pp. 187-190.

[7] D. Socolinsky, and A. Selinger, "Face recognition in the dark," Conf. on Computer Vision and Pattern Recognition Workshop, 2004, vol. 18, no. 8, pp. 129-134.

[8] F. Prokoski, "History, current status, and future of infrared identification," Proceedings of IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications, pp. 5-14, 2000.

[9] MK. Bhowmik, K. Saha, S. Majumder, G. Majumder, A. Saha, AN. Sarma, D. Bhattacharjee, K. Dipak Basu, and M. Nasipuri. Thermal infrared face recognition – a biometric identification technique for robust security system, Reviews, Refinements and New Ideas in Face Recognition, edited by Peter M. Corcoran, ISBN: 978-953-307-368-2, InTech, 2011.

[10] AM. Guzman, M. Goryawala, J. Wang, A. Barreto, J. Andrian, N. Rishe, and M. Adjouadi, "Thermal imaging as a biometrics approach to facial signature authentication," IEEE JNL Biomed Health Inform, vol. 17, no. 1, pp. 214-222, 2013.

[11] Papakostas, G.A. ; Kaburlasos, V.G. ; Pachidis, T., "Thermal infrared face recognition based on lattice computing (LC) techniques,", IEEE International Conference on Fuzzy Systems, pp. 1 – 6, 2013.

[12] Seal, A. ; Nasipuri, M. ; Bhattacharjee, D. ; Basu, D.K., "Thermal human face recognition based on GappyPCA," IEEE Second International Conference on Image Information Processing (ICIIP), pp. 597 – 600, 2013.

[13] S. Li, L. Zhang, S. Liao, X. Zhu, R. Chu, M. Ao, and R. He, "A near-infrared image based face recognition system," In Proc. IEEE International Conference on Automatic Face and Gesture Recognition, pages 455–460, 2006.

[14] S. Li, R. Chu, S. Liao, and L. Zhang, "Illumination invariant face recognition using near-infrared images," IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(4):627–639, 2007.

[15] Khairul Hamimah Abas and Osamu Ono, "Implementation of Frontal-Centroid Moment Invariants in Thermal-Based Face Identification System," Fifth International Conference on Signal Image Technology and Internet Based Systems, 2009.

[16] S. Wu, W. Song, L.J. Jiang, S. Xie, F. Pan, W.Y. Yau, and S. Ranganath, "Infrared face recognition by using blood perfusion data," In Proc. International Conference on Audio- and Video-Based Biometric Person Authentication, pages 320–328, 2005.

[17] Zhihua Xie, Guodong Liu and Shiqian Wu , Zhijun Fang, "Infrared Face Recognition Based on Blood Perfusion and Fisher Linear Discrimination Analysis," IST 2009 - International Workshop on Imaging Systems and Techniques, Shenzhen, China, May 11-12, 2009.

[18] J. Flusser and T. Suk, "Rotation moment invariants for recognition of symmetric objects," IEEE Transactions on Image Processing, vol. 15, no. 12, pp. 3784–90, 2006.

[19] Rota Bulò, S., Albarelli, A., Torsello, A. and Pelillo, M., "A hypergraph-based approach to affine parameters estimation," in Proceedings of the 19th International Conference on Pattern Recognition ICPR'08 (Tampa, Florida), 2008.

[20] Suk, T. and Flusser, J., "Affine moment invariants generated by automated solution of the equations," in Proceedings of the 19th International Conference on Pattern Recognition ICPR'08 (Tampa, Florida), 2008.