

Implementation of Multi-Centroid Moment Invariants in Thermal-Based Face Identification System

Khairul Hamimah Abas and Osamu Ono
Department of Electronics and Bioinformatics,
School of Science and Engineering, Meiji University, Kanagawa, Japan

Abstract: Problem statement: Current paramount methods and approaches in face identification field rely on facial characteristics; such as location of eyes, length of nose and mouth, regardless the type of medium used to acquire the facial images. The visibility of these facial characteristics varies significantly with environmental factors (e.g., lighting elements). Various researches have been devoted to develop methodologies for addressing these problems. Despite the overwhelming effort to overcome these problems, the current framework of face identification research has a conceivable weakness due to its very nature. **Approach:** In this study, we presented a novel framework for face identification based on thermal information extracted from facial images acquired within the thermal spectrum. The motivation initiating to conduct this research was to engage in extracting significant facial characteristics from the acquired thermal information, which differs from current facial characteristics that are visible over the skin. Primarily, we proposed a non-holistic analysis approach by decomposing the input image into several input images via multiple threshold method (3 threshold values). This allowed us to analyze the thermal information given within a specific thermal range. Then, the first moment invariant, I_1 , from Hu's classical moment invariants were computed and extracted (with respect to centroid point obtained from frontal shot source images) which constituted the feature database. For each subject, we used one (frontal shot) registered image with two and four test images (left, mid left, right and mid right). This approach employs minimum distance measurement method for classification purposes. **Results:** We had conducted experiments on the IRIS IR facial dataset with encouraging results of 95% of correct identification rate for test to registered image ratio of 2:1 and 92.5% of correct identification rate for test to registered image ratio of 4:1. **Conclusion:** The obtained results demonstrated the high capability of Hu's classical moment invariants as a feature in thermal based face identification and introduce new ways for classical methods to further be utilized in theoretical and empirical research in the area.

Key words: Biometrics, face identification, infrared, moment invariants, centroid

INTRODUCTION

As humans, we have always relied on holistic facial identification as a mean to identify another person's identity. Due to this remarkable capability that exists in humans, face identification-based security systems is gaining major acceptance as paramount biometric approach in access control and surveillance. It is non-invasive, hygiene and natural since it coexists with the mode of identification that humans employ in everyday matters. Although humans are capable of identifying faces effortlessly, this similar task is trivial in security vision system. Analysis and identification of facial images which was acquired from a real and non-ideal imaging system still holds many complications

since the appearance of face in concern varies dramatically due to incident angle and light variation, facial expressions, head pose and image quality.

Most research efforts in face identification field had focused on visible spectrum imagery (Kong *et al.*, 2005). Even though progress in visible spectrum imagery is notably high, problems due to the nature of the approach used in visible spectrum is still an open issue. Images acquired from visible spectrum are formed primarily due to reflection and due to this; it is difficult to process because of the high dependency on incident angle and lighting variation from external light source. Recently, interest in thermal infrared imagery had increased among researchers. Thermal Infrared (IR) imagery is basically based on heat emission. Since

Corresponding Author: Khairul Hamimah Abas, Department of Electronics and Bioinformatics,
School of Science and Engineering, Meiji University, Kanagawa, Japan

thermal IR imagery is independent to external light source, problems encountered in visible spectrum-based systems do not exist in thermal IR imagery. This was proven by works done in Socolinsky *et al.* (2006) and Friedrich and Yeshurun (2002).

A more complex work was done in (Buddharaju *et al.*, 2007) which utilize the physiological features. The obtained physiological features are closely related to the distribution of blood vessels under the face skin. The features extracted are used to form a unique thermal faceprint. Distance transform is then used to obtain an invariant representation for face identification. This work tremendously achieved 86% in correct identification rate for the University of Houston database. Geometric moment invariants were utilized for object identification (Rizon *et al.*, 2006). Canny and Sobel edge detector was employed for segmentation prior to the calculation of geometric moment invariants. The first four of Hu's moment invariants were extracted as features from each image. This study (Rizon *et al.*, 2006) achieved 90% in correct identification rate with Sobel edge detector, while, by employing the Canny edge detector approach, it could only achieved up to 70% in correct identification rate. Despite the maximum correctness in identification rate for both aforementioned works, works done in (Buddharaju *et al.*, 2007) implements complex algorithms. This would further complicate the realization process in security vision system.

Motivated by works done in (Buddharaju *et al.*, 2007) and also inspired by the implementation of Hu's classical moment invariants for object classification done in (Rizon *et al.*, 2006), this study presents the implementation of moment invariants (with respect to centroid point obtained from frontal mugshot images) for thermal-based face identification system. Generally, most face identification system practices holistic analysis approach. This differs from the proposed approach where analysis is done non-holistically within each decomposed thermal regions. Likewise, the propose approach offers a broader possibilities for its usage. Due to independency towards facial features; such as length of eyes, nose and mouth, the proposed approach has the potential to be implemented for identification of other body parts; such as arms, abdomens and legs. The proposed system initially filters the background scenery via seeded region growing method. Later, it decomposes a filtered infrared image into 4 thermal regions via 3-valued threshold method. Nevertheless, region with the lowest thermal value is omitted from further processes. This is to ensure that the background scenery is not taken into

consideration. Thereafter, Hu's first moment invariants (with respect to centroid point obtained from frontal view of each individual) is calculated as means of feature to be extracted from each generated thermal region. Minimum distance measurement is employed for classification between Hu's first moment invariants for stored and tests images. This is similar to template matching method. Further throughout this study, it can be seen that the proposed approach did not involve any training phase.

The outline of this study is as follow: The methods demonstrate the application of our proposed approach, followed by empirical results and discussion. Sequentially, conclusions are drawn. Finally, references are presented at the end of this study.

MATERIALS AND METHODS

Acquired facial image normally contains background scenery. If the entire image is taken into consideration for feature extraction, it may affect the performance of the system. Therefore, we employed seeded region growing method to remove the background scenery. This is done alongside with other conventional image pre-processes; such as histogram equalization and image normalization, prior to the proposed method.

The following subsections demonstrate the application of 3-valued threshold for thermal region decomposition which is followed by centroid point calculation (for later use in Hu's classical moment invariants), elaboration on Hu's classical moment invariants and also the implementation of minimum distance measurement for classification.

3-valued threshold decomposition: The general definition of threshold is represented by the following equation:

$$g(x,y) = \begin{cases} 255, & f(x,y) \geq T \\ 0, & f(x,y) < T \end{cases} \quad (1)$$

Where:

$f(x,y)$ = The input pixel

$g(x,y)$ = The output pixel

T = The threshold value

By inserting three threshold values rather than one threshold value, the 3-valued threshold equation can be derived from (1) as follows:

$$g(x,y) = \begin{cases} L_4, & f(x,y) \geq T_3 \\ L_3, & T_2 < f(x,y) \leq T_3 \\ L_2, & T_1 < f(x,y) \leq T_2 \\ L_1, & f(x,y) \leq T_1 \end{cases} \quad (2)$$

Where:

T_1, T_2, T_3 = The three threshold values
 L_1, L_2, L_3, L_4 = The label for each generated thermal regions

This shows that an image can be decomposed into several thermal regions by substituting the threshold values with border line thermal values. As a result, four thermal regions (four binary formatted images) are generated. In order to obtain these thermal regions, values for T_1, T_2 and T_3 is selected based on the results acquired from the preliminary experiments conducted in (Abas *et al.*, 2009). Referring to works done in (Abas *et al.*, 2009), an initial value for T_1, T_2 and T_3 is randomly selected within the range stated in (Buddharaju *et al.*, 2004) where the temperatures at all pixels are mapped between 0 and 255. Mapped temperatures between 200 and 225 is said to be common temperature on face and mapped temperature between 175-200 and 225-255 are said to be normal temperature on cheeks and maximum temperature on face, respectively. Since the area for maximum mapped temperature on face is small and sparsely located within a face, this would cause the system to identify these areas as noise. Therefore, we selected T_3 to have the initial value of 200; the minimum value for the combination of mapped temperatures for common and maximum temperature on face.

As aforementioned, mapped temperatures between 175 and 200 is said to be normal temperatures on cheeks. With manual tuning done in (Abas *et al.*, 2009), we discovered that the mapped temperatures between approximately 140 and 200 comprehend temperatures on convex surfaces of a face; such as nose, cheeks and forehead. Hence, the initial value for T_2 is set to 140. By employing the same manual tuning technique used in (Abas *et al.*, 2009), T_1 is initially set to approximately 80, where this value affirms with values stated in (Buddharaju *et al.*, 2004) (mapped temperature value between 0 and 100 normally indicates the background scenery). For maximum assurance that the background scenery is not taken into consideration, the lowest valued region (coldest region) is omitted from further processes. Although we have proposed the usage of multiple thresholds (constitute of 3 threshold values), various number of threshold values (e.g., 4, 5... N threshold values) could also be implemented, thus producing more thermal segments for further analysis.

Centroid point: Referring to an online source (Wikipedia, 2009), the first moment invariant, Φ_1 (which will be elaborated in the next subsection), is

roughly proportional to the moment of inertia around the image's centroid, if the pixel's intensities were interpreted as physical density. Thus the following is a brief interpretation on centroid calculation.

The general definition for centroid point in a discrete mass is as follows:

$$C = \frac{\sum m_i r_i}{\sum m_i} \quad (3)$$

where, r_i and m_i are particle positions and mass, respectively. For a binary formatted image, the numerator's particle position, r_i , is substituted with pixel's coordinate and mass, m_i , is substituted with pixel's intensity. For the denominator, $\sum m_i$ is substituted with the total number of pixels with the intensity of 1. The derived equation is as follows:

$$C(x) = \frac{\sum I(x_i)}{\sum \text{pixel}} \quad (4)$$

$$C(y) = \frac{\sum I(y_i)}{\sum \text{pixel}} \quad (5)$$

where, $C(x)$ and $C(y)$ are coordinates for x-axis and y-axis, respectively. Therefore, the actual location of centroid point for a binary formatted image becomes $(C(x), C(y))$. Note that, intensity, I , holds a value of 1 or 0 for binary formatted images and 0-255 for a gray-scaled image.

Moment invariants: Image moments are weighted averages (moments) of the image pixels' intensities, or functions of those moments, usually chosen to have some attractive property or interpretation. These moments are normally used in numerous fields; such as image processing and computer vision. They are useful to describe objects after segmentation. Simple properties of the image which are found via image moments include area (or total intensity), centroid and information about its orientation.

Similarly, moment invariants are properties of connected regions in images that are invariant to translation, rotation and scale. They are useful because they define a simple calculated set of region properties that can be used for shape classification and pattern recognition. Invariants to similarity transformation; such as rotation, translation and scaling, was the first invariants that appeared in pattern recognition field. It was caused partly because of their simplicity, partly because of great demand for invariant features that

could be used in position-independent object classification. In this problem formulation, degradation operator, D , is supposed to act solely in spatial domain and to have a form of similarity transform. Initially, the relation between the ideal image $f(x,y)$ and the observed image $g(x,y)$ is described as $g = D(f)$. Due to invariants to similarity transformation, this results to the following equation:

$$g(r(x,y)) = f(x,y) \quad (6)$$

where, $r(x,y)$ denotes arbitrary rotation, translation and scaling.

Invariants to translation and scaling are trivial in any imaging system. As early as, Hu (1962) formulated the famous seven rotation invariants, consisting second and third order moments. Followings are examples of the first four of Hu's moment invariants:

$$\begin{aligned} \Phi_1 &= \mu_{20} + \mu_{02} \\ \Phi_2 &= (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2 \\ \Phi_3 &= (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \\ \Phi_4 &= (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2 \end{aligned} \quad (7)$$

As aforesaid, the first moment invariant, Φ_1 , is roughly proportional to the moment of inertia around the image's centroid, if the pixel's intensities were interpreted as physical density. By interpreting the pixel's intensities as physical density, we calculated the first moment invariant for each decomposed thermal region (with respect to centroid point acquired from each background filtered, frontal view source images). This is shown in Fig. 1. Theoretically, the impact of each thermal region (from different input image for the same individual) surrounding one common centroid will be similar; or slightly deviates, if the source images are similar or slightly differs in angular pose. Therefore, the moments acquired from each respected thermal region surrounding the images' centroid will result to similar values. Followings are the derivation of Hu's first moment invariant:

$$\phi_1 = \eta_{20} + \eta_{02} \quad (8)$$

Where:

$$\begin{aligned} \eta_{20} &= \frac{\mu_{20}}{\mu_{00}^2} \\ \eta_{02} &= \frac{\mu_{02}}{\mu_{00}^2} \end{aligned} \quad (9)$$

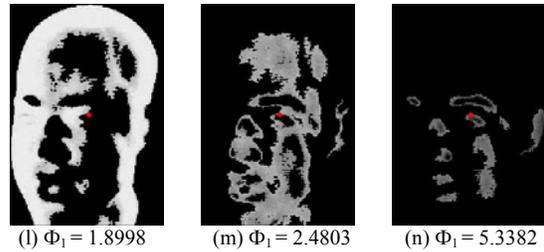
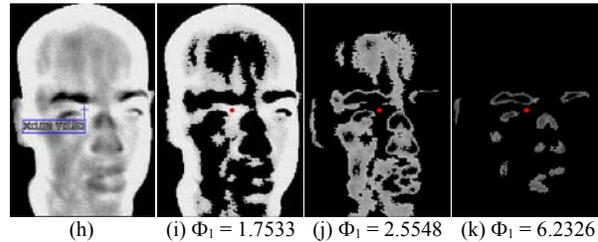
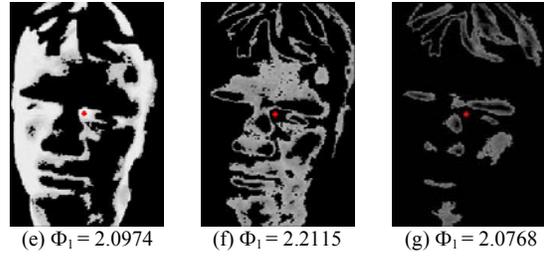
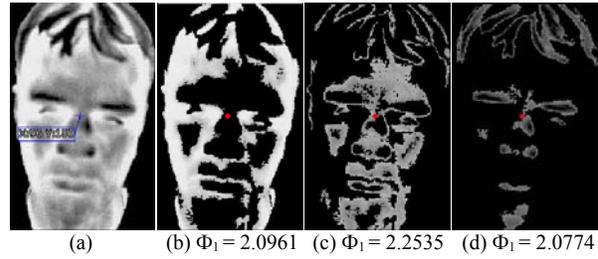


Fig. 1: Example of Hu's first moment invariants calculated values. (a) and (h): Frontal source image; cross-hair represents the centroid point used as references for Hu's first moment invariant calculation. (Second to fourth column from left): Examples of decomposed thermal regions generated via 3-valued threshold. The red dot represents the centroid point obtained from each individual frontal view image ((a) and (h))

and:

$$\begin{aligned} \mu_{20} &= M_{20} - \bar{x}(M_{10}) \\ \mu_{02} &= M_{02} - \bar{y}(M_{01}) \\ \mu_{00} &= M_{00} \end{aligned} \quad (10)$$

where \bar{x} and \bar{y} are centroid point coordinate for x-axis and y-axis, respectively, while raw image moment, M_{ij} with pixel intensity, $I(x,y)$ are calculated as follows:

$$M_{ij} = \sum_x \sum_y x^i y^j I(x,y) \quad (11)$$

for $i,j = 0, 1$ and 2 .

On the other hand, the centroid point is obtained from the holistic frontal view of registered images prior to the calculation of moment invariant (or moment inertia) for each decomposed thermal region.

Minimum distance measurement: Previously, Hu's first moment invariants were calculated and extracted as features to be classified. Hence, for classification, we employed minimum distance measurement method between the stored and test values of Hu's first moment invariants obtained from each corresponding thermal region. The general definition of minimum distance measurement, x , via Euclidean Distance between two points, P and Q, is shown in (12):

$$x = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2} \quad (12)$$

For higher order of points, the distance between any finite number of points, n , is shown in (13) below:

$$x = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (13)$$

In our case, since only two values are being compared; therefore (12) is redefined as follows:

$$x = \sqrt{(r - t)^2} \quad (14)$$

where, r and t are values of Hu's first moment invariants for registered and test images, respectively.

RESULTS

OTCBVS IRIS IR facial database were used to validate the effectiveness of our proposed approach. Figure 2 shows some examples from this dataset. All calculations were done with MATLAB 7.0 Student Version on a 1.8 GHz Centrino Duo processor with 1 GB RAM.

We conducted two experiments on the OTCBVS IRIS IR facial database to evaluate the performance of the proposed face identification method. We have used 2 and 4 test images for 1 registered image. Each image is decomposed into 4 thermal regions where the lowest

(coldest) thermal region is not taken into consideration. This results to a total of 6 and 12 decomposed test images used for 3 decomposed registered images per individual. We would like to point out that only frontal shot images were used for registered images; whereas images with slight angular deviation in pose; to the left and right were used for test images. The first set of test images (2 test images) consists of quarter-left and quarter-right profiles, whereas the second set of test images (4 test images) constitute of quarter-left, mid-left, quarter-right and mid-right profiles. Since only the frontal shot images were used for registered dataset, left and right profiles were not included in the test dataset.

In the first experiment, we implemented the original Hu's first moment invariant, Φ_1 , for face identification of both sets of 2 and 4 test images. In the second experiment, we employed the proposed approach; the first moment invariant (with respect to centroid point obtained from frontal shot source images), for both sets of test images. The performance for both experiments is shown in Fig. 3. Figure 3a shows the Cumulative Match Characteristic (CMC) curves for the first experiment and Fig. 3b shows the CMC curves for the second experiment. We have compared the identification performance of our approach with works done in (Buddharaju *et al.*, 2007). The comparison of performance between these two approaches is shown in Fig. 4.



Fig. 2: Samples of OTCBVS IRIS IR facial dataset

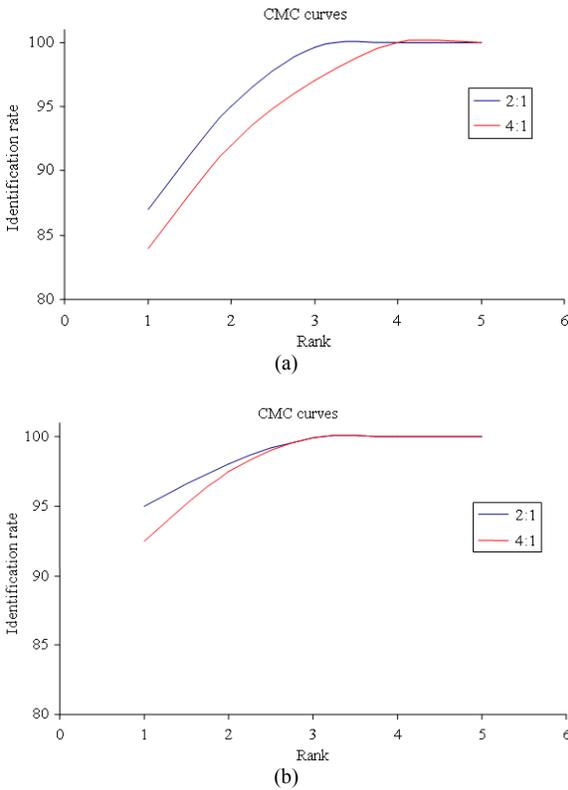


Fig. 3: CMC Curves of (a) the original Hu's moment invariant for test to registered image ratio of 2:1 and 4:1 and (b) our approach for test to registered image ratio of 2:1 and 4:1

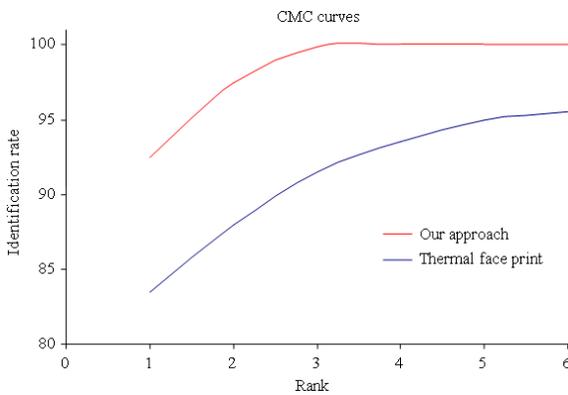


Fig. 4: CMC curves of our approach and works done in (Buddharaju *et al.*, 2007)

DISCUSSION

Referring to Fig. 3a and b, experiment conducted on the first set of 2 test images performs better than the

second set of 4 test images. This is due to test cases that are close to mid-left and mid-right profiles in the second set of 4 test images may be less accurate, since only frontal shot images are being used as registered images.

Referring to Fig. 4, the CMC curves shows that rank 1 identification for our approach is 92.5%, which exceeds the performance for work done in (Buddharaju *et al.*, 2007) (approximately 83.5%). Albeit our approach demonstrates encouraging performance, the robustness of this approach degrades when health issues are being addressed (fever and flu). At the moment, this matter is considered as the operational limit for this approach.

CONCLUSION

In this study, we had demonstrated a novel approach in infrared image processing. Like many other approaches, conventional pre-processes were employed prior to our approach. For maximum assurance, we employed seeded region growing method for background filtering. A 3-valued threshold were derived and employed for region decomposition. We also introduced the implementation of Hu's moment invariants in multiple thermal regions as a feature to be extracted within each generated thermal regions. Our approach was tested on OTCBVS IRIS IR facial database which is publicly available for download at www.cse.ohio-state.edu/otcbvs-bench/. Classifications are done by employing a minimum distance measurement method between the acquired moment invariant from test and registered IR images. Empirical results obtained show that with proper enhancement of classical methods, the performance of our proposed approach surpasses the performance of works done by Buddharaju *et al.* (2007) with some operational limitations.

ACKNOWLEDGMENT

The researchers would like to thank Prof. Marzuki Khalid and especially Dr. Zuwairie Ibrahim for the time and guidance given throughout the preliminary experiments, University of Technology, Malaysia and SLAB for giving the opportunity to pursue PhD in Meiji University and also to the OTCBVS IRIS IR facial Database for the images used in the experiments.

REFERENCES

- Abas, K.H., O. Ono, K. Suzuki and S. Yoshida, 2009. Thermal segmented-based face identification system. Proceeding of the Asia Simulation Conference, Oct. 7-9, Japan Society for Simulation Technology (JSST), Japan, pp: ID007. <http://www.keepandshare.com/doc/1779655/id007-pdf-march-2-2010-4-05-pm-189k?da=y>
- Buddharaju, P., I. Pavlidis and I. Kakadiaris, 2004. Face recognition in the thermal infrared spectrum. Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition Workshop, June 27-July 2, IEEE Computer Society, USA., pp: 133. DOI: 10.1109/CVPR.2004.343
- Buddharaju, P., I.T. Pavlidis, P. Tsiamyrtzis and M. Bazakos, 2007. Physiology-based face recognition in the thermal infrared spectrum. *IEEE Trans. Patt. Anal. Machine Intell.*, 29: 613-626. DOI: 10.1109/TPAMI.2007.1007
- Friedrich, G. and Y. Yeshurun, 2002. Seeing people in the dark: Face recognition in infrared images. Proceedings of the Second International Workshop on Biologically Motivated Computer Vision, Nov. 22-24, Springer-Verlag, London, pp: 348-359. <http://portal.acm.org/citation.cfm?id=751741>
- Hu, M.K., 1962. Visual pattern recognition by moment invariants. *IRE Trans. Inform. Theor.*, 8: 179-187. <http://ieeexplore.ieee.org/iel5/4547527/22787/01057692.pdf>
- Kong, S.G., J. Heo, B.R. Abidi, J.Paik and M.A. Abidi, 2005. Recent advances in visual and infrared face recognition-a review. *J. Comput. Vis. Image Understand.*, 97: 103-135. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.106.9877&rep=rep1&type=pdf>
- Rizon, M., H. Yazid, P. Saad, A.Y.M. Shakaff and A.R., Saad *et al.*, 2006. Object detection using geometric invariant moment. *Am. J. Applied Sci.*, 3: 1876-1878. <http://www.scipub.org/fulltext/ajas/ajas361876-1878.pdf>
- Socolinsky, D.A., L.B. Wolffz, J.D. Neuheisely and C.K. Evelandz, 2006. Illumination invariant face recognition using thermal infrared imagery. Report Number A763444, 2006, Equinox Corporation, New York. <http://www.equinoxsensors.com/publications/face.pdf>
- Wikipedia, 2009. Image moments. Wikipedia. http://en.wikipedia.org/wiki/Image_moments