

Lower-Dimensional Feature Sets for Template-Based Motion Recognition Approaches

Md. Atiqur Rahman Ahad, J. Tan, H. Kim and S. Ishikawa
Department of Control Engineering, Kyushu Institute of Technology, Fukuoka, Japan

Abstract: Problem statement: In template-based motion recognition approaches, feature sets are computed from the template for classification. Hu invariants are widely employed for this purpose since its inception. However, development of lower-dimensional feature vector sets is required for faster computation along with robust recognition. The concept of reduced size of Hu moment is really interesting. From its inception, seven higher orders Hu moments have been employed by many researchers without considering why seven and why not less numbers. **Approach:** In this study, we analyzed with various feature sets with different number of Hu moments and rationalized that based on the characteristics of central moments, it is not necessary to employ all the seven moments in every applications and, in that way, we can reduce the computational cost and make it faster. **Results:** Based on various feature vectors sets, it is evident that we can use lower dimensional feature vectors for our Directional Motion History Image (DMHI) method and other methods. **Conclusion:** Therefore, we can conclude that we do not need all seven invariants, rather 1st two or three invariants seem enough-as we are not reproducing the image. Higher invariants are noisy and hence can be ignored. The 0th order moment for Energy images provide enough information about the mass area and hence, no need to calculate the other seven invariants.

Key words: Hu moment, DMHI, feature vector, motion recognition

INTRODUCTION

Action understanding and motion recognition has various important applications and therefore, development of robust action classification methods are crying needs in the computer vision arena. For classification and action recognition, moment invariants are widely employed (Ahad *et al.*, 2008a) to develop smart feature vector sets. Moment invariants were firstly introduced to the pattern recognition community by (Hu, 1962), who employed the results of the theory of algebraic invariants and derived his seven famous invariants to rotation of 2D objects. Since its inception, the Hu (1962) invariants became classical. The use of moments for image analysis, object representation and recognition was inspired by these invariants. However, both for image reconstruction and pattern recognition, researchers employ all the seven invariants in their study for shape analysis. Even though it is assumed that higher order moments are less stable, detailed experimental analysis has been unattended and hence, all the invariants are computed. In general, gross image shape is represented well by the lower-order moments and high-order moments only reflect the subtleties of a silhouette or boundary image. We find that complexity

increases dramatically with increasing order and their containment of redundant information about shape.

We estimated these invariants to calculate feature vectors using our Directional Motion History Image (DMHI) method (Ahad *et al.*, 2008b; 2009). We employed seven moments to one moment, with and without normalized 0th order moments, to create 12 different feature vectors for classification. It seems that 0th order moment is not required to consider, though for energy images, it can give key information about the mass of the motion area and hence, we can ignore to employ seven moments for energy images. However, instead of using seven moments for energy images, we can consider only the 0th order moments that provide the total object area, as it seems sufficient and, in this way, we can reduce feature vectors. Therefore, these feature vectors can significantly reduce the computational cost without sacrificing the recognition rates. We vividly notice that the average recognition rates are still in satisfactory ranges even with lower-dimensional feature vectors. We also present various graphical demonstrations by which it becomes clear that the higher order moments are less useful for recognition purposes. It is evident that higher the invariants, more the variations in them. Higher the

order for invariants, more noisier and unstable natures they demonstrate. We note that the absolute error rate is much higher for higher order moments compared to the lower order moments. Therefore, for pattern recognition purpose, lower order moments are the key to achieve more sound recognition.

The concept of reduced size of Hu moment is really interesting. From its inception, 7 higher orders Hu moments have been employed by many researchers without considering why 7, why not less numbers. We analyzed with various feature sets with different number of Hu moments and rationalized that based on the characteristics of central moments, it is not necessary to employ all the seven moments in every applications and, in that way, we can reduce the computational cost and make it faster. Therefore, we can conclude that we do not need all 7 invariants, rather 1st 2/3 invariants seem enough-as we are not reproducing the image. Higher invariants are noisy and hence can be ignored. The 0th order moment for Energy images provide enough information about the mass area and hence, no need to calculate the other seven invariants.

Background: As pointed by Ahad *et al.* (2008a); Moeslund *et al.* (2006); Poppe (2007) and Aggarwal and Cai (1999), human activity and motion recognition has various paradigms and various approaches are considered for pattern recognition. Hu moments are widely used by many researchers for image representation or image reconstruction, utilized as pattern recognition features in a number of applications, such as, action recognition (Ahmad and Lee, 2006; Bradski and Davis, 2002; Bobick and Davis, 2001; Dudani *et al.*, 1977), for fingerprint verification (Yang and Park, 2007), texture classification (Campisi *et al.*, 2004; Nacereddine and Tridi, 2005), rapid matching of video streams, data matching (Wong and Hall, 2002), character recognition (El-Khaly and Sid-Ahmed, 1990; Tsirikolias and Mertzios, 1993), image normalization (Gruber and Hsu, 1997) and estimation of position and the attitude of the object in 3-D space (Mukundan and Ramakrishnan, 1996).

In general, gross image shape is represented well by the lower-order moments and high-order moments only reflect the subtleties of a silhouette or boundary image (Mukundan and Ramakrishnan, 1996). Hence, one may not even need all seven moment invariant functions to design a classifier (Mukundan and Ramakrishnan, 1996).

Prokop and Reeves (1992) noted in their survey on moment-based techniques that most practical experiments have shown little improvement in

identification performance when moment orders are increased beyond order 4 or 5 and in general, high order moments are very sensitive to noise and less stable (Celebi and Aslandogan, 2005; Prokop and Reeves, 1992; Shen and Ip, 1999; Teh and Chin, 1998).

In another research, for shape analysis and classification of weld defects in industrial radiography, only the first two Hu invariants are employed along with some other geometric parameters (Nacereddine and Tridi, 2005). The first two invariants give measures in relation with the pixel spreading in comparison with the center of mass (Nacereddine and Tridi, 2005). The first two moment invariants were used by Hu (1962) to represent several known digitized patterns in a two-dimensional feature space. In another research, Rizon *et al.* (2006) used only the first invariants for object detection. They claimed that from 2nd invariant onward, the invariants are insignificant.

Bidoggia and Gentili (2002) pointed that it is possible to see that the 4th and the 5th moments have the larger standard deviations (especially for R-G and B-Y channels in color images) and that the moments are less stable under rotation and scaling. They also commented that this is not surprising, because the higher order moments are the less stable. In order to maintain a good stability, a smaller set of moments has been selected, composed by the five moments in gray-scale image (and for color image, by three moments in both R-G and B-Y images). The higher is the order of the moment, the higher are the fluctuations (Bidoggia and Gentili, 2002). As all the complex moments are approximate, the last moments are the less stable, because they depend on higher power of uncertain numbers. One study by (Teh and Chin, 1998) showed that higher order moments are more vulnerable to white noise, thus making their use undesirable for pattern recognition with higher order moments.

However, these invariants have several drawbacks. Information redundancy is one of the drawbacks. Since the basis is not orthogonal, these moments suffer from a high degree of information redundancy (Celebi and Aslandogan, 2005). Moreover, in the presence of noise, the computed Hu invariant moments, begin to degrade. Also, large variation in the dynamic range of values may create instability. Since the basis involves powers of p and q , the moments computed have large variation in the dynamic range of values for different orders. This may cause numerical instability when the image size is large (Celebi and Aslandogan, 2005). The concept of lower-dimensional feature vector sets and the relevance of it for action recognition by employing our directional motion history image method is presented below. Moreover, we employ other methods for further analysis.

MATERIALS AND METHODS

Since its inception, seven Hu moments are widely used. Similarly, we employ Hu moments to develop feature vectors for the Directional Motion History Image (DMHI) representations for each activity. Here we compute the (DMHI) templates based on a threshold ξ on pixel values:

$$\begin{aligned}
 H_t^{-x}(x,y,t) &= \begin{cases} \tau & \text{if } \Psi_x^-(x,y,t) > \xi \\ \max(0, H_t^{-x}(x,y,t-1)-\delta) & \text{otherwise} \end{cases} \\
 H_t^{+x}(x,y,t) &= \begin{cases} \tau & \text{if } \Psi_x^+(x,y,t) > \xi \\ \max(0, H_t^{+x}(x,y,t-1)-\delta) & \text{otherwise} \end{cases} \\
 H_t^{-y}(x,y,t) &= \begin{cases} \tau & \text{if } \Psi_y^-(x,y,t) > \xi \\ \max(0, H_t^{-y}(x,y,t-1)-\delta) & \text{otherwise} \end{cases} \\
 H_t^{+y}(x,y,t) &= \begin{cases} \tau & \text{if } \Psi_y^+(x,y,t) > \xi \\ \max(0, H_t^{+y}(x,y,t-1)-\delta) & \text{otherwise} \end{cases}
 \end{aligned}$$

Here:

- (x,y) = Pixel position
- t = Temporal duration
- δ = The decay factor
- $\Psi(x,y,t)$ = The optical flow vector

In this equation, this vector is split into four different channels according to horizontal and vertical directions:

$$\mathbb{W} \in \left\{ \text{up} \left(\begin{matrix} +y \\ \end{matrix} \right), \text{down} \left(\begin{matrix} -y \\ \end{matrix} \right), \text{right} \left(\begin{matrix} +x \\ \end{matrix} \right), \text{left} \left(\begin{matrix} -x \\ \end{matrix} \right) \right\}$$

From these DMHI templates, we get its binaries energy image templates called DMEI. In this study, we employ these templates.

Lets define the 2D (p+q)th order Cartesian moment m_{pq} of a density distribution function $\rho(x,y)$:

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q \rho(x,y) dx dy \quad (p,q=0,1,2,\dots)$$

where, p and q are the order of the moments in the x and y axes, respectively. A complete moment set of order n consists of all moments, m_{pq} such that $p+q \leq n$ and contains $\frac{1}{2}(n+1)(n+2)$ elements. The moment transformation for a 2D digital image $f(x,y)$, of size (N,M) is:

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

Hu derived relative and absolute combinations of moment values that are invariant with respect to scale, position and orientation based on the theories of invariant algebra that deal with the properties of certain classes of algebraic expressions which remain invariant under general linear transformations. These moments are not invariant to geometrical transformations. To achieve invariance under translation, we need to get central moment. The central moment μ_{pq} can be defined as:

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q \rho(x,y) d(x - \bar{x}) d(y - \bar{y})$$

Here, $\bar{x} = m_{10} / m_{00}$, $\bar{y} = m_{01} / m_{00}$. This is essentially a translated Cartesian moment, which means that the centralized moments are invariant under translation. The first four orders (i.e., (p+q) is from 0-3) are defined as:

$$\begin{aligned}
 \mu_{00} &= m_{00} \equiv \mu \\
 \mu_{10} &= \mu_{01} = 0 \\
 \mu_{20} &= m_{20} - \mu \bar{x}^2 \\
 \mu_{11} &= m_{11} - \mu \bar{x} \bar{y} \\
 \mu_{02} &= m_{02} - \mu \bar{y}^2 \\
 \mu_{30} &= m_{30} - 3m_{20} \bar{x} + 2\mu \bar{x}^3 \\
 \mu_{21} &= m_{21} - m_{20} \bar{y} - 2m_{11} \bar{x} + 2\mu \bar{x}^2 \bar{y} \\
 \mu_{12} &= m_{12} - m_{02} \bar{x} - 2m_{11} \bar{y} + 2\mu \bar{x} \bar{y}^2 \\
 \mu_{03} &= m_{03} - 3m_{02} \bar{y} + 2\mu \bar{y}^3
 \end{aligned}$$

For the second and third order moments, following seven orthogonal invariants are achieved to calculate feature vectors. The first six moments are rotation, scaling and translation invariants. 7th moment is skew (and bi-correlations) invariant that enables it to distinguish mirror or otherwise identical images. The seven invariants are:

$$\begin{aligned}
 I_1 &= \mu_{20} + \mu_{02} \\
 I_2 &= (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2 \\
 I_3 &= (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \\
 I_4 &= (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2 \\
 I_5 &= (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] \\
 &\quad + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \\
 I_6 &= (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^3] \\
 &\quad + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) \\
 I_7 &= (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 \\
 &\quad - 3(\mu_{21} + \mu_{03})^2] - (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03}) \\
 &\quad [3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]
 \end{aligned}$$

Based on these seven invariants, for the DMHI and DMEI templates, we need to compute seven Hu moments for each of the eight components. Moreover, to cover the overwriting issue significantly, we introduce the normalized 0th order Hu moments ($\overline{m_{00}^{\varpi}}$). The 0th order moments provide the total object area or mass and hence provide important cues for motion region. Finally, considering seven invariants for each along with eight normalized 0th order moment invariants ($\overline{m_{00}^{\varpi}}$ where $\varpi \in \{\text{up, down, left, right}\}$ -directing four directions), we compute a 64-D feature vector, FV_{dmhi} (i.e., 7 Hu moments for each of the DMHI and DMEI template (i.e., $7 \times 8 = 56$), in addition to eight (1×8) normalized 0th order moments. Therefore, we get a feature vector of dimensions of $56+8 = 64$). Let us calculate 0th order moment for template $\text{med}H_{\tau}^{+x}(x, y, t)$ as:

$$m_{00}(H_{\tau}^{+x}) = \sum_x \sum_y \text{med}H_{\tau}^{+x}(x, y, t)$$

Then we get normalized 0th order moment for $\text{med}H_{\tau}^{+x}(x, y, t)$ by employing the following equation:

$$\overline{m_{00}^{+x}} = \frac{m_{00}(H_{\tau}^{+x})}{m_{00}(H_{\tau}^{+x}) + m_{00}(H_{\tau}^{-x}) + m_{00}(H_{\tau}^{+y}) + m_{00}(H_{\tau}^{-y})}$$

These normalized 0th order invariants confer the relative mass of the motion area in the scene. We can define the final feature vector, FV_{dmhi} as:

$$FV_{\text{dmhi}} = (\forall h)(I_{1 \rightarrow 7}) \cup (\forall e)(I_{1 \rightarrow 7}) \cup (\forall h, \forall \varpi) \left(\overline{m_{00}^{\varpi}} \right) \cup (\forall e, \forall \varpi) \left(\overline{m_{00}^{\varpi}} \right)$$

where, h denotes the motion history image components and e stands for energy components of the DMHI method. However, we notice that for Hu moment, complexity increases dramatically with increasing order and their containment of redundant information about shape. The normalized 0th order moment provides a ration of the area. It is a measure of compactness.

RESULTS

We employ seven moments to one moment, with and without normalized 0th order moments, to create 12 different feature vectors for classification (Table 1). We also consider moments only for history templates and only for energy templates and achieved different

recognition results. It seems that 0th order moment is not required to consider, though for energy images, it can give key information about the mass of the motion area and hence, we can ignore to employ 7 moments for energy images. However, instead of using seven moments for energy images, we can consider only the 0th order moments that provide the total object area, as it seems sufficient and, in this way, we can reduce from 64-36D feature vectors.

For recognition purpose, we employ our dataset of various aerobics. Ten different actions (e.g., Body stretching; Waving arms, bending and straighten legs; Turning the arms; Bending the chest; Bending the body sideway (left); Bending the body-front and back; Waving arms and twisting the body to the left; Bending and straightening arms up to shoulders, legs move; Bending the body diagonally to bottom-both sides and Bending the body sideway (right)) from eight various subjects are taken from an uncelebrated frontal-view digital video camera with almost constant illumination condition in indoor environment. No special markers or dress or arrangement were considered in this experiment. Also the dress, height, size and age of the subjects were different for each person. The frame has resolution of 320×240 pixels. For classification, we consider leave-one-out cross-validation approach. In Table 1, we present some comparative recognition results for the DMHI method with various feature vector sets (Ahad *et al.*, 2008c). In this process, we employ k-nearest neighbor classification method for recognition based on our database. In this analysis, we tried with seven moments to one moment to make 12 feature vector sets for classification (Ahad *et al.*, 2008c). It seems that 0th order moment is not required to consider, though for energy images, it can give key information about the mass of the motion area and hence, we can ignore to employ 7 moments for energy images.

Table 1: Recognition rate with various feature vectors employing DMHI and DMEI

FV set	Hu moments		Rec. rate (%)
	For DMHI	For DMEI	
64D	7 + norm. 0th order	7 + norm. 0th order	94.4
56D	7 Hu moments	7 Hu moments	89.0
48D	6 Hu moments	6 Hu moments	89.0
40D	5 Hu moments	5 Hu moments	89.0
36D	7 + norm. 0th order	norm. 0th order	97.3
32D0th	7 Hu moments	norm. 0th order	89.0
32D	4 Hu moments	4 Hu moments	89.0
24D	3 Hu moments	3 Hu moments	89.0
16D	2 Hu moments	2 Hu moments	89.0
8D	1 Hu moments	1 Hu moments	89.0
4D	1 Hu moments	0	80.6
4D	0	1 Hu moments	44.5

If we consider 7 or 6 or 5 or 4 moments to calculate feature vectors for each history and energy template, we achieve the same recognition rate. We evaluated the top five results for all recognitions and we did not find any change in its distribution, as the distances among these are almost the same even if we consider 7 or 4 moments. However, if we use the first 3 Hu moments, then we notice changes among the results even though the recognition result is not varying significantly. Finally, we tried with history images only and energy images only to get recognition result and it shows poor recognition as it is evident that for better recognition, we need to consider both motion history and energy templates. Therefore, reduced numbers of Hu moments are suitable and well-fit for this DMHI method. In this way, we can reduce the running time.

Some of these different feature vector sets can be shown as (every sub-equation here denotes feature vectors of 56, 40, 36, 32 and 24D respectively):

$$\begin{aligned} \text{FV}_{56\text{D}} &= (\forall h)(I_{1 \rightarrow 7}) \cup (\forall e)(I_{1 \rightarrow 7}) \\ \text{FV}_{40\text{D}} &= (\forall h)(I_{1 \rightarrow 5}) \cup (\forall e)(I_{1 \rightarrow 5}) \\ \text{FV}_{36\text{D}} &= (\forall h)(I_{1 \rightarrow 7}) \cup (\forall h)(\overline{m_{00}^0}) \cup (\forall e)(\overline{m_{00}^0}) \\ \text{FV}_{32\text{D}} &= (\forall h)(I_{1 \rightarrow 7}) \cup (\forall e)(\overline{m_{00}^0}) \\ \text{FV}_{24\text{D}} &= (\forall h)(I_{1 \rightarrow 3}) \cup (\forall e)(I_{1 \rightarrow 3}) \end{aligned}$$

For example, $\text{FV}_{56\text{D}}$ consists of seven invariants for both history and energy templates, whereas, $\text{FV}_{32\text{D}}$ set is composed with seven invariants for history images along with the normalized 0th order moments of the energy images. Therefore, these feature vectors can significantly reduce the computational cost without sacrificing the recognition rates much.

We can vividly notice that the average recognition rates are still in satisfactory ranges even with lower-dimensional feature vectors. We find it a wonderful development with the DMHI method, because in all other approaches that have employed Hu invariants in their works, have calculated the seven invariants, even though higher invariants are more noise-prone. In the case with the DMHI, the deployment of lower dimensional feature vectors can also produce good recognition results and we achieved average recognition rates of 91%. Therefore, we can find a faster recognition strategy with the DMHI approach considering lower-dimensional feature sets. In the similar fashion, we have analyzed with the Hierarchical Motion History Histogram (HMHH) method (Meng *et al.*, 2006). We consider four patterns in this case and computed with various feature sets. We found that until the initial three invariants, the

recognition rates are almost unchangeable. Similarly, by employing the basic Motion History Image (MHI) method by Bobick and Davis (2001), we achieve similar conclusions. Therefore, we can conclude that for recognition purpose, it is not required to employ all invariants-rather initial few invariants seem enough for template-based motion recognition methods.

DISCUSSION

Here present various graphical demonstrations by which it becomes clear that the higher order moments are less useful for recognition purposes. It is evident that higher the invariants, more the variations in them. Figure 1 shows the variations of the 64 invariants (i.e., seven Hu invariants and normalized 0th order moments for each of the four history images and four energy images). In Table 2 and Fig. 1-5, we use 'fv' to denote 'feature vector number'.

Table 2 shows the definition of the different features; e.g., fv1~fv7 denotes the seven invariants for $H_{\tau}^{+x}(x, y, t)$ where fv1 is for the 1st invariant, fv2 denotes the 2nd invariant and so on. Figure 2 shows the variations of the seven different Hu invariants in log-scale for $H_{\tau}^{+x}(x, y, t)$ (i.e., fv1~fv7). From this Fig. 2, it is clearly understood that the first invariant is stable compared to its higher invariants. Higher the order for invariants, more noisier and unstable natures they demonstrate.

Table 2: Various Feature Vectors (FV)'s definition

FV	Definition/comment
fv1~fv7	1st seven invariants for $H_{\tau}^{+x}(x, y, t)$
fv8~fv14	1st seven invariants for $H_{\tau}^{-x}(x, y, t)$
fv15~fv21	1st 7 invariants for $H_{\tau}^{+y}(x, y, t)$
fv22~fv28	1st seven invariants for $H_{\tau}^{-y}(x, y, t)$
fv29~fv35	1st seven invariants for $E_{\tau}^{+x}(x, y, t)$
fv36~fv42	1st seven invariants for $E_{\tau}^{-x}(x, y, t)$
fv43~fv49	1st seven invariants for $E_{\tau}^{+y}(x, y, t)$
fv50~fv56	1st seven invariants for $E_{\tau}^{-y}(x, y, t)$
fv57~fv60	Norm. 0th order moments for $H_{\tau}^{+x}(x, y, t)$, $H_{\tau}^{-x}(x, y, t)$, $H_{\tau}^{+y}(x, y, t)$ and $H_{\tau}^{-y}(x, y, t)$
fv61~fv64	Norm. 0th order moments for $E_{\tau}^{+x}(x, y, t)$, $E_{\tau}^{-x}(x, y, t)$, $E_{\tau}^{+y}(x, y, t)$ and $E_{\tau}^{-y}(x, y, t)$

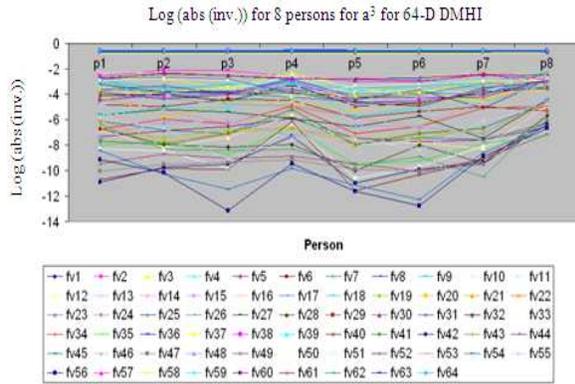


Fig. 1: All 64 invariants of the DMHI for various persons for an action

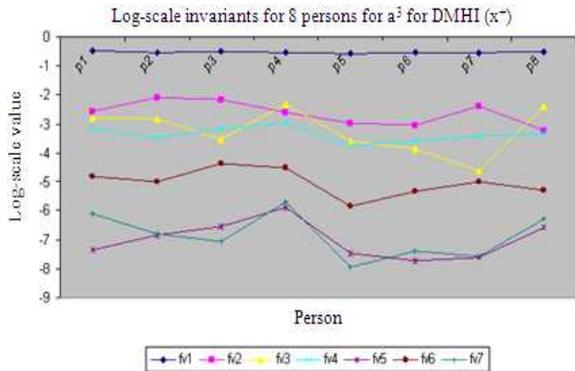


Fig. 2: Invariants variations for $H_{\tau}^{+x}(x, y, t)$ -these clearly shows that higher the invariants, more the variations

Figure 3 shows the normalized 0th order moments (i.e., fv57~fv64) and the 1st invariant for all eight directional-history and directional-energy templates (i.e., fv1, fv8, fv15 and fv22 are for four DMHI templates and fv29, fv36, fv43 and fv50 show the corresponding 1st invariants for four DMEI templates respectively). From Fig. 4, graph for the 1st invariants are presented. We see that the range for y-axis is 0-0.9 only and the ranges of values are from -0.77 to -0.47 (for 1st invariants). So the variation is less. For 6th and 7th invariants, we will see wide range of values.

In the similar fashion, we compared the 6th and 7th invariants for the same action for eight different persons. Figure 5 demonstrates these for 6th invariants (i.e., fv6, fv13, fv20, fv27, fv34, fv41, fv48 and fv55 for four DMHIs and four DMEI respectively) and for 7th invariants (i.e., fv7, fv14, fv21, fv28, fv35, fv42, fv49 and fv56 for four DMHIs and four DMEI respectively). It is clearly seen that the y-axis is very large (0-14.0) compared to that of Fig. 5 (0-0.9).

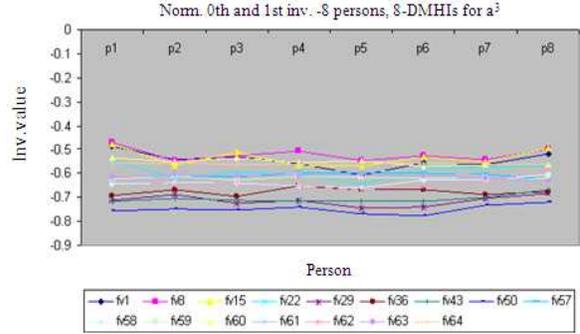


Fig. 3: Normalized 0th order and 1st invariants for 8 persons' action

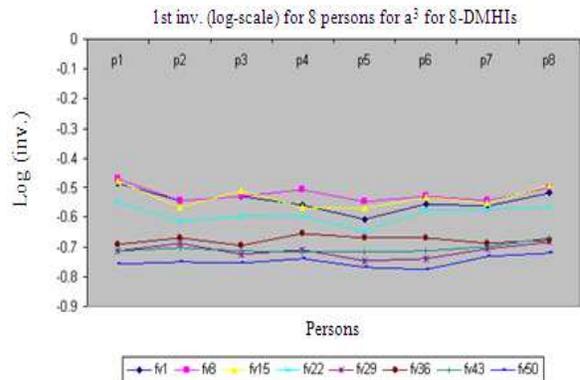


Fig. 4: For 8 different persons' action, 1st invariants is shown. Range for y-axis is 0-0.9

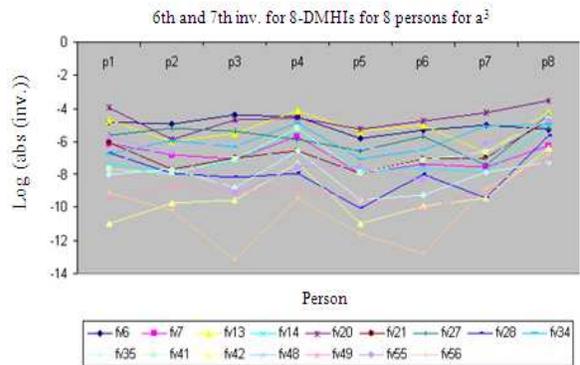


Fig. 5: 6th and 7th invariants for 8 different persons' action (for both energy and history images). Y-axis is larger (0~14) than the same for 0th and 1st invariants (0~0.9)

Also the variations among the same invariant for different persons are high and hence less stable. The range for the DMEIs is from -4.4-13.15 and the same for

DMHIs is from-3.55-10.0. These features show clearly that higher invariants are more unstable and noisy.

We also analyze the absolute error for some invariants to see the variation of the invariants from the first person to the rest of the subjects, which demonstrate that the higher order moments are noisy and unstable. Therefore, for pattern recognition purpose, lower order moments are the key to achieve more sound recognition.

CONCLUSION

Since its inception, the Hu's invariants became classical and, despite of their few drawbacks, numerous works have been devoted to various application areas. The use of moments for image analysis and object representation was inspired by these invariants. We notice that most of the cases, both for image reconstruction and pattern recognition, researchers employ all the seven invariants in their study for shape analysis. Even though it is assumed that higher order moments are less stable, detailed experimental analysis has been unattended and hence, all the invariants are computed.

We analyzed with various feature sets with different number of Hu moments and rationalized that based on the characteristics of central moments, it is not necessary to employ all the seven moments in every applications and, in that way, we can reduce the computational cost and make it faster. Based on various FV sets, it is evident that we can use lower dimensional feature vectors for the DMHI and other methods. HMHH and MHI methods are also exploited and we achieved similar results. Therefore, we can conclude that we do not need all seven invariants, rather 1st two or three invariants seem enough-as we are not reproducing the image. Higher invariants are noisy and hence can be ignored. The 0th order moment for Energy images provide enough information about the mass area and hence, no need to calculate the other seven invariants.

ACKNOWLEDGMENT

Researchers are thankful to the students who participated in creating the action database.

REFERENCES

Aggarwal, J.K. and Q. Cai, 1999. Human motion analysis: A review. *Comput. Vision Image Understand.*, 73: 428-440. DOI: 10.1006/cviu.1998.0744

Ahad, M.A.R., J. Tan, H. Kim and S. Ishikawa, 2009. Temporal motion recognition and segmentation approach. *Int. J. Image Syst. Technol.*, 19: 91-99. DOI: 10.1002/ima.v19:2

Ahad, M.A.R., J. Tan, H. Kim and S. Ishikawa, 2008a. Human activity recognition: Various paradigms. *Proceeding of the International Conference Control Automation and Systems*, Oct. 14-17, IEEE Xplore Press, Seoul, Korea, pp: 1896-1901. DOI: 10.1109/ICCAS.2008.4694407

Ahad, M.A.R., T. Ogata, J. Tan, H. Kim and S. Ishikawa, 2008b. A complex motion recognition technique employing directional motion templates. *Int. J. Innovat. Comput. Inform. Control*, 4: 1943-1954.

Ahad, M.A.R., J. Tan, H. Kim and S. Ishikawa, 2008c. Template-based human motion recognition for complex activities. *Proceeding of the IEEE International Conference on Systems, Man and Cybernetics*, Oct. 12-15, IEEE Xplore Press, Singapore, pp: 673-678. DOI: 10.1109/ICSMC.2008.4811355

Ahmad, M. and S.W. Lee, 2006. Human action recognition using multi-view image sequences features. *Proceeding of the International Conference on Automatic Face and Gesture Recognition*, Apr. 2-6, IEEE Xplore Press, Southampton, UK., pp: 523-528. DOI: 10.1109/FGR.2006.65

Bidoggia, R. and S. Gentili, 2002. A basis of invariant moments for color images. *Proceeding of the International Workshop on Systems, Signals and Image Processing*, Nov. 7-8, World Scientific, Singapore, pp: 527-531. DOI: 10.1142/9789812776266_0080

Bobick, A. and J. Davis, 2001. The recognition of human movement using temporal templates. *IEEE Trans. PAMI*, 23: 257-267. DOI: 10.1109/34.910878

Bradski, G. and J. Davis, 2002. Motion segmentation and pose recognition with motion history gradients. *Mach. Vis. Appli.*, 13: 174-184. DOI: 10.1007/s001380100064

Campisi, P., A. Neri, G. Panci and G. Scarano, 2004. Robust rotation-invariant texture classification using a model-based approach. *IEEE Trans. Image Process.*, 13: 782-791. DOI: 10.1109/TIP.2003.822607

Celebi, M. and Y. Aslandogan, 2005. A comparative study of three moment-based shape descriptors. *Proceeding of the International Conference on Information Technology, Coding and Computing*, Apr. 4-6, IEEE Xplore Press, USA., pp: 788-793. DOI: 10.1109/ITCC.2005.3

- Dudani, S., K. Breeding and R. McGhee, 1977. Aircraft identification by moment invariants. *IEEE Trans. Comput.*, 26: 39-45. DOI: 10.1109/TC.1977.5009272
- El-Khaly, F. and M. Sid-Ahmed, 1990. Machine recognition of optically captured machine printed Arabic text. *Patt. Recogn.*, 23: 1207-1214. DOI: 10.1016/0031-3203(90)90116-3
- Gruber, M. and K.Y. Hsu, 1997. Moment-based image normalization with high noise-tolerance. *IEEE Trans. PAMI*, 19: 136-139. DOI: 10.1109/34.574793
- Hu, M.K., 1962. Visual pattern recognition by moment invariants. *IRE Trans. Inform. Theory*, IT-8: 179-187. DOI: 10.1109/TIT.1962.1057692
- Meng, H., N. Pears and C. Bailey, 2006. Recognizing human actions based on motion information and SVM. *Proceeding of the IEE International Conference Intelligent Environments*, July 5-6, Athens, Greece, pp: 239-245. DOI: 10.1049/cp:20060648
- Moeslund, T.B., A. Hilton and V. Kruger, 2006. A survey of advances in vision-based human motion capture and analysis. *Comput. Vis. Image Understand.*, 104: 90-126. DOI: 10.1016/j.cviu.2006.08.002
- Mukundan, R. and K. Ramakrishnan, 1996. An iterative solution for object pose parameters using image moments. *Patt. Recogn. Lett.*, 17: 1279-1284. DOI: 10.1016/0167-8655(96)00099-2
- Nacereddine, N. and M. Tridi, 2005. Computer-aided shape analysis and classification of weld defects in industrial radiography based invariant attributes and neural networks. *Proceeding of the International Symposium on Image and Signal Processing and Analysis, (ISPA'05)*, Zagreb, Croatia, pp: 88-93. DOI: 10.1109/ISPA.2005.195389
- Poppe, R., 2007. Vision-based human motion analysis: An overview. *Comput. Vis. Image Understand.*, 108: 4-18. DOI: 10.1016/j.cviu.2006.10.016
- Prokop, R. and A. Reeves, 1992. A survey of moment-based techniques for unoccluded object representation and recognition. *CVGIP Graph. Models Image Process.*, 54: 438-460. DOI: 10.1016/1049-9652(92)90027-U
- Rizon, M., H. Yazid, P. Saad, A. Shakaff and A. Saad *et al.*, 2006. Object detection using geometric invariant moment. *Am. J. Applied Sci.*, 2: 1876-1878.
- Shen, D. and H.S.H. Ip, 1999. Discriminative wavelet shape descriptors for recognition of 2-D patterns. *Patt. Recogn.*, 32: 151-165. DOI: doi:10.1016/S0031-3203(98)00137-X
- Teh, C. and R. Chin, 1998. On image analysis by the methods of moments. *IEEE Trans. PAMI*, 10: 496-513. DOI: 10.1109/34.3913
- Tsirikolias, K. and B.G. Mertzios, 1993. Statistical pattern recognition using efficient two-dimensional moments with applications to character recognition. *Patt. Recogn.*, 26: 877-882. DOI: 10.1016/0031-3203(93)90053-Y
- Wong, R. and E. Hall, 2002. Scene matching with invariant moments. *Comput. Graph. Image Process.*, 8: 16-24. DOI: 10.1016/S0146-664X(78)80028-8
- Yang, J., J.W. Shin and D. Park, 2007. Fingerprint matching using invariant moment features. *Lecturer Notes Comput. Sci.*, 4465: 1029-1038. DOI: 10.1007/978-3-540-74377-4