NEW HUMAN ACTION RECOGNITION SCHEME WITH GEOMETRICAL FEATURE REPRESENTATION AND INVARIANT DISCRETIZATION FOR VIDEO SURVEILLANCE

NILAM NUR BINTI AMIR SJARIF

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor Philosophy (Computer Science)

Faculty of Computing
Universiti Teknologi Malaysia

MARCH 2015

To my beloved father, Amir Sjarif Hj Chairan, and my mother, Zaima Abu Bakar, and my lovely sister's, Syarfawati, Sofea Iman and Nur Fatihah.

ACKNOWLEDGMENT

In the name of Allah, the Most Gracious and the Most Merciful. I thank to Allah for granting me strength and guidance throughout my journey to complete my study.

My grateful thanks to Prof. Dr. Siti Mariyam Shamsuddin and Assoc. Prof. Dr. Siti Zaiton Mohd Hashim, for their help, guidance and encouragement throughout the duration of my studies. I would like to thank Ministry of Higher Education (MOHE) under Fundamental Research Grant Scheme (Vote 4F347) that supports my research work. I am also very thankful to Soft Computing Research Group (SCRG) and UTM Big Data Centre for their guidance, advices and motivation.

Special thanks to all my colleagues especially my lovely sister's Kak Intan Ermahani, Kak Shafaatunnur, Kak See Pheng, Kak Yee Leng, Kak Shalimatunnur, Kak Umi, Kak Chachada, Syazwa, Farhana Hordri and Asilah for their support. Special thanks also to my lovely best friends Nor Syaliza, Zanariah, Nur Fazrina and Dzidatul Akma. Without them, I would not this far. Thank you so much for the encouragement and insightful conversation over the year. My sincere appreciation also extends to all others who have provided assistance at various occasions. Their views and tips are useful indeed.

Finally, to my beloved Parents (Amir Sjarif Bin Hj Chairan and Zaima Binti Abu Bakar) and my lovely sweet sister's (Syarfawati, Sofea Iman and Nur Fatihah), thank you so much for their support, love and encouragement throughout my life. They are with me in every way. Thank you so much for everything.

ABSTRACT

Human action recognition is an active research area in computer vision because of its immense application in the field of video surveillance, video retrieval, security systems, video indexing and human computer interaction. Action recognition is classified as the time varying feature data generated by human under different viewpoint that aims to build mapping between dynamic image information and semantic understanding. Although a great deal of progress has been made in recognition of human actions during last two decades, few proposed approaches in literature are reported. This leads to a need for much research works to be conducted in addressing on going challenges leading to developing more efficient approaches to solve human action recognition. Feature extraction is the main tasks in action recognition that represents the core of any action recognition procedure. The process of feature extraction involves transforming the input data that describe the shape of a segmented silhouette of a moving person into the set of represented features of action poses. In video surveillance, global moment invariant based on Geometrical Moment Invariant (GMI) is widely used in human action recognition. However, there are many drawbacks of GMI such that it lack of granular interpretation of the invariants relative to the shape. Consequently, the representation of features has not been standardized. Hence, this study proposes a new scheme of human action recognition (HAR) with geometrical moment invariants for feature extraction and supervised invariant discretization in identifying actions uniqueness in video sequencing. The proposed scheme is tested using IXMAS dataset in video sequence that has non rigid nature of human poses that resulting from drastic illumination changes, changing in pose and erratic motion patterns. The invarianceness of the proposed scheme is validated based on the intra-class and inter-class analysis. The result of the proposed scheme yields better performance in action recognition compared to the conventional scheme with an average of more than 99% accuracy while preserving the shape of the human actions in video images.

ABSTRAK

Pengecaman aksi manusia merupakan bidang penyelidikan yang aktif dalam visi komputer kerana terdapatnya lambakan aplikasi yang sering digunakan dalam bidang pengawasan video, capaian semula video, sistem keselamatan, pengindeksan video dan interaksi komputer manusia. Pengecaman aksi dikelaskan sebagai perubahan ciri data yang dihasilkan oleh manusia dari sudut pandangan berbeza adalah bertujuan untuk membina pemetaan antara maklumat imej yang dinamik dengan pemahaman semantik. Walaupun terdapat banyak kemajuan yang telah dibuat terhadap pengecaman aksi manusia dalam dua dekad yang lalu, beberapa pendekatan yang dicadangkan dalam kajian kesusasteraan dilaporkan. Ini membawa kepada keperluan untuk melaksanakan kerja-kerja penyelidikan bagi menangani cabaran yang berterusan dan membangunkan pendekatan yang lebih cekap bagi menyelesaikan masalah pengecaman aksi manusia. Pengekstrakan ciri merupakan tugasan utama dalam pengecaman aksi yang mewakili teras dalam prosedur pengecaman aksi. Dalam pengawasan video, momen takubahan global berdasarkan Momen Takubah Geometri (GMI) lazimnya digunakan dalam pengecaman aksi. Walau bagaimanapun, terdapat banyak kelemahan menggunakan GMI seperti kekurangan tafsiran relatif tidak varian secara terperinci terhadap bentuk. Oleh yang demikian, perwakilan ciri tidak terselaraskan. Maka, kajian ini mencadangkan satu skema baru bagi pengecaman aksi manusia dengan menggunakan fungsi GMI untuk pengekstrakan ciri dan seliaan pendiskretan bagi mengenalpasti keunikan aksi manusia dalam jujukan video. Skema cadangan diuji dengan menggunakan dataset IXMAS bagi imej-imej dalam jujukan video yang mempunyai sifat tidak tegar postur manusia yang terhasil akibat daripada perubahan drastik pencahayaan, perubahan postur dan corak pergerakan yang tidak menentu. Ketakubahan skema cadangan disahkan berdasarkan analisis kelas-intra dan kelas-inter. Hasil kajian menunjukkan bahawa prestasi bagi skema cadangan adalah lebih baik dalam pengecaman aksi berbanding dengan skema tradisi iaitu dengan kadar pengecaman adalah lebih daripada 99% ketepatan dengan mengekalkan bentuk aksi manusia dalam imej video.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

$I(x, y, t)$	-	The density of the volume pixel at time t
τ	-	The number of frames in action cycle
dfa	-	Degree of freedom associated with SSB
dga	-	Degree of freedom associated with SSW
n	-	The number of image frame
x_i	-	The current image frame
r_i	-	The reference image frame (first image frame is the reference image)
f	-	The number of features image frame
i	-	The feature's column of image frame
α	-	Significance level
$H_0; H_1$	-	ANOVA hypothesis
m_{pq}	-	Moment order (p, q)
$f(x, y)$	-	Density function
$F - value$	-	Parameter to reject or aspect the hypothesis H_0 in ANOVA
L	-	The number of pixels present in the image frames
μ_{pq}	-	Center moment
η_{pq}	-	Scaling factor
$\theta_1 - \theta_7$	-	Invariant feature of GMI order i
$\boldsymbol{v_{min}}$	-	Minimum value of invariant feature vector for human action
$\boldsymbol{v_{max}}$	-	Maximum value of invariant feature vector for human action
$\boldsymbol{iv}$	-	Interval value
$\boldsymbol{iv_{min}}$	-	Minimum value of an interval for human action
$\boldsymbol{iv_{max}}$	-	Maximum value of an interval for human action
$\boldsymbol{rv}$	-	Representation value

LIST OF ABBREVIATIONS

2D	-	Two-dimensional
2GSS	-	Second Generation Surveillance System
3D	-	Three-dimensional
3GSS	-	Third Generation Surveillance System
AMI	-	Aspect Moment Invariant
ANOVA	-	Analysis Variance
AUMI	-	Aspect United Moment Invariant
BF	-	Best First
BGE	-	Between-Groups estimate of variance
BMI	-	Bamieh Moment Invariant
C4.5	-	J48
CCTV	-	Closed circuit television
CMI	-	Complex Moment Invariant
CMU	-	Carnegie Mellon University
CVPR	-	Computer Vision & Pattern Recognition
DARPA	-	Defense Advanced Research Projection Agency
EFB	-	Equal Frequency Binning
EM	-	Expectation Maximization
EWB	-	Equal Width Binning
FD	-	Fuzzy Discretization
FT	-	Functional Trees
GMI	-	Geometric Moment Invariant
HAI	-	Human action invarianceness
HAR	-	Human Action Recognition
HARS	-	Human Action Recognition Scheme
HCI	-	Human computer interaction

HGMI	-	Higher Order Scale Moment Invariant
HID	-	Human Identification at a Distance
HMA	-	Human Motion Analysis Model
HUMI	-	Higher United Moment Invariant
IFD	-	integrated feasibility demonstration
IGSS	-	First Video Surveillance System
IRBD	-	Information Retrieval Binning Discretization
IXMAS	-	Inria Xmas Motion Acquisition Sequences
LMI	-	Legendre Moment Invariant
MAE	-	Mean absolute error
MF	-	Moment functions
NPZMI	-	Normalized Pseudo Zernike Moment Invariant
NS	-	Not Significant
PCA	-	Principal Component Analysis
PZMI	-	Pseudo Zernike Moment Invariant
RMI	-	Regular Moment Invariant
S	-	Significant
SNK	-	Student Newman-Keul
SSB	-	Sum of square between group
SSW	-	Sum of square within group
TMI	-	Tchebichef Moment Invariant
TZMI	-	Teague Zernike Moment Invariant
UMI	-	United Moment Invariant
VSAM	-	Visual Surveillance and Monitoring
VSS	-	Video Surveillance System
WCMI	-	Weighted Central Moment Invariant
WEKA	-	Waikato Environment for Knowledge Analysis
WGE	-	Within-Groups estimate of variance
WMI	-	Weighted Moment Invariant
ZMI	-	Zernike Moment Invariant

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Overview

Of late, computer vision and pattern recognition have gained a lot of interest among researchers, motivated by the promise of the various applications in different domains. The applications include video surveillance, human computer interaction, video editing, gesture recognition and a wide range of industries such as robot learning and control, sports and biometric.

One of the popular areas in computer vision and pattern recognition is human motion analysis, specifically in human action recognition (HAR). The aim of human action recognition is to generalize the variation of the human actions, and through the process, the image in various invariant positions can be featured. However human actions recognition remains a challenging task due to the dependency of the feature representation and accurate recognition (Barnachon *et al*., 2014; Bloom *et al*., 2013; Li *et al*., 2014; Cheng *et al*., 2014; Donderac *et al.*, 2009; Junejo *et al*., 2013; Lara and Labrador, 2013; Wu *et al*., 2011; Weinland *et al.*, 2007; Weinland *et al*., 2006; Parameswaran and Chellappa, 2005).

1.2 Background of the Problem

Human action recognition (HAR) has become essential parts of many computer vision applications, and it involves a series of processes such as image data acquisition, feature extraction and representation and classification for recognition (Kaiqi *et al.*, 2012; Samy Sadek *et al.*, 2013). A human action can be viewed as a set of sequential silhouettes, whereby the basic approach of human action recognition is to extract a set of features from each frame of sequence (Di and Ling, 2013), then the features is used to perform classification (Rao and Shah, 2001). A human silhouette contains the detailed information of the shape of the body, which can be obtained using translation, rotation and scaling processes (Di and Ling, 2013; Kui and Dit-Yan, 2008; Gorelick, Galun *et al,* 2006; Ahmad, Mohiuddin Lee, *et al*., 2008; Chaaraoui, *et al*., 2013) .

The advancement of human silhouette is commonly used as action description. Examples of human silhouette approach includes the history of shape changes using 2D action temporal templates (Ahad, *et al*., 2012; Bobick and Davis, 1996; Bobick and Davis, 2001; Lu, *et al*., 2012; Megavannan, *et al*., 2012), and the extension of 2D templates to 3D action template (Weinland, *et al.*, 2006). Similarly, the notations of action cylinders and space time shapes were introduced based on silhouettes by Gorelick, *et al*. (2006), and Wang and Leckie (2010). All the works mentioned above rely mostly on the effective feature extraction and a representation technique, which is normally, combined with machine learning or pattern recognition methods.

In the past two decades, many action recognition techniques have been developed based on feature extraction and representation (Zhao, *et al*., 2013). Feature extraction and representation, a fundamental part of action recognition, greatly influences the performance of the recognition system. The representation of features are divided into two categories: local representation and global representation (Calderara, *et al.*,2008; Ikizler, *et al*.,2008; Mokhber, *et al*.,2008; Mokhber, *et al*.,2005; X.Sun, *et al*.,2009; Zhao, *et al*., 2013). Local representation is described as a collection of independent patches which features are extracted from the region of interest. The limitations of the local feature include the sparse representation that could discard geometric relationship of the features, hence is less discriminative (Di and Ling, 2013; Poppe, 2010). Nevertheless, the representation is dependent on the extraction of sufficient amount of relevant interest points, which focused more on local motion than the figure shape, and sometimes pre-processing is needed to compensate for camera movements. Meanwhile, global representations are powerful since much of the information is encoded, as they focus on global information. Common global representations are derived from silhouettes, edges or optical flow.

Based on the literatures, many studies have been done on global approach for human action recognition. Bobick and Davis (2001), Bradski and Davis (2000), Davis and Bradski (1999), Hota,Venkoparao, *et al*. (2007) and Rosales (1998) were employed global representation for feature extraction and representation. In their works, global representation gives reasonable shape discrimination in translation and scaled invariant manner for template matching. Therefore, the process of extraction can be less computation and feasible implementation in real-time by using global based approach. Achard, *et al*. (2008), Mokhber, *et al*. (2005), Parameswaran and Chellappa (2005) were successfully extract the features based on global representation, so all information concerning an action is included in only one vector and allows to recognize actions without systems such as finite state machines. Therefore, the recognition process can be simplify and enhance robustness. Global representations statistically matched to stored examples of different movements and give a promising result using a large database of movements (Davis and Bradski, 1999). Thus, the representations of sequence are not characterized as the temporal object (Achard, *et al*., 2008; X.Sun, *et al*., 2009). And through the previous study by Ahmad, *et al.* (2008) has been identified that the global representation is very useful moments kernel and present a native rotational invariance and far more robust to noise.

The conventional global representation based moment invariant that was put forward by Hu (1962), Geometric Moment Invariant (GMI), was widely used for feature extraction and classification in human action recognition (Megavannan, *et* *al*., 2012). This method has a unique characteristic in identifying an image due to its invariant to orientation, size and position of the shape image. The existence of action is identified by difference values of moment invariants in different image frame in video sequence. Consequently, this method leads some issues, especially in terms of the complexity of data feature representation during the feature extraction. The problem examples are in the intra-class and inter-class variation. The complexity increases dramatically with increasing order and their containment of redundant information about shape. Many researchers have found some drawbacks of the GMI, especially in pattern recognition. Table 1.1 illustrates the research developments in solving the weaknesses of GMI.

Table 1.1 Drawbacks of Geometric Moment Invariant Concept

Problem/ Issues	Researchers/Year
Reliant and incomplete invariant under translation, rotation and scaling	(Flusser and Suk, 1993; Xu and Li, 2008; Zhihu and Jinsong, 2010)
Lost scale invariant in discrete condition	(Botao Wong, *et al*., 2002; Ding, *et al*., 1992; Hongtao and Jicheng, 1993; Lihong, *et al*., 2006)
Improved boundary images condition	(Chen, 1993)
Applied only a small subset of moment invariant	(Wong, *et al*., 1995)
Produce errors if the transformations are subjected to unequal scaling data transformation	(Feng and Keane, 1994; Raveendran, *et al*., 1993; Muda, *et al*., 2007; Muda, *et al*.,2008; Palaniappan, *et al*., 2000; Pamungkas and Shamsuddin, 2009; Raveendran, *et al*., 1994; Raveendran, *et al*., 1995a,b; Raveendran, *et al*., 1997; Shamsuddin, *et al*., 2000; Shamsuddin, *et al*., 2002)
Data position of pixel is far away from centre coordinate	(Balslev, *et al*., 2000)
Problem in region, boundary and discrete condition	(Yinan, *et al*., 2003)

A study conducted in 1992 has identified that the scale invariant was lost when the condition is in discrete condition (Ding, *et al*., 1992). Consequently, new techniques based on moment variant in discrete condition have been proposed by various researchers such as Hongtao and Jicheng (1993) and Botao Wong, *et al*. (2002), and a technique known as relative contour moment invariant was also introduced by Lihong, *et al.* (2006). Also, some drawbacks has been identified in

Hu moments that the approach are reliant on whereby incomplete invariant occurs under translation, rotation and scaling (Flusser and Suk, 1993). Thus, a new moment formulations called affine moment invariant was presented by Flusser and Suk (1993), which is an invariant under general affined transformation. This technique has been successfully used in human action recognition for feature extraction in Samy Sadek, *et al.* (2013). In 2008, the issue regarding the extracting shape characteristics independent of scaling, translation and rotation is discussed by Xu and Li (2008). Thus, for a good solution to this problem, Xu and Li presented 3-D moment invariants which is invariants under similarity transformation and the moments are region-based.

In 1995, a group of researchers have identified that GMI can only be applied on a small subset of moment invariants when they found that GMI is not able to determine a complete set of Moment Invariant (Wong, *et al*., 1995). Thus, they improved the third-order moment and fourth-order moment of GMI and tested the methods on Character recognition. In 1993, a derivation on normalization for the improvement of the conditions of boundary images, called Improved Moment Invariant, has been proposed by Chen (1993). The advantages of the proposed method is that it could save computational time for the boundary images. However, it was discovered that the techniques mentioned earlier cannot be applied to both region and boundary conditions as the features represented by the equations are not the same as that obtained from the GMI methods (Yinan, *et al.*, 2003). Consequently, they proposed a new technique for translation, scaling and rotations that can be discretely kept invariant into region and boundary conditions. A study conducted in year 2000 has identified that GMI have some problem with regards to the position of data, in which the pixel of data is far away from the centre of coordinate if there is a noise (Balslev, *et al.*, 2000). This consequence happened because GMI cannot effectively recognize the data that are concentrated near the center of mass, and as a result these data will be abandoned.

Another drawback of GMI is problems associated with transformation errors. GMI produce error if the transformations are subjected to unequal scaling. Based on the literature review conducted, numerous researchers have made various improvements on the invariants for images that undergo unequal scaling such as Pamungkas and Shamsuddin (2009), Muda, *et al.* (2008), Muda, *et al.* (2007), Shamsuddin, *et al.* (2002), Shamsuddin, *et al*. (2000), Palaniappan, *et al.* (2000); Feng and Keane (1994), Raveendran, *et al*. (1993), Raveendran, *et al*. (1997), Raveendran, *et al.* (1995a) Raveendran, *et al*. (1995b), Raveendran, *et al*. (1994). Normally, the scaling factor used in GMI is equal scaling, but it does not work well for images that undergo unequal scaling. Unequal scaling means that the x and y-scaling of an image are different (Feng, and Keane, 1994). Table 1.2 illustrates the advancement of unequal scaling formulation based on the normalization centre to obtain invarianceness when unequal scaling is used.

Table 1.2 Advancement of Geometrical Moment in Unequal Scaling Images

Authors/Year	Description
Raveendran, *et al.* (1993)	This equation has a flaw of scaling in the derivation of the central moments. Mathematical expression: $\eta_{pq} = \gamma_{pq}\gamma_{qp}\ ;\ \gamma_{pq} = \left(\frac{b}{a}\right)^{\frac{p-q}{2}}$
Raveendran, *et al*. (1994)	Mathematical expression: $\tilde{\lambda}_{pq} = r^{\frac{p-q}{2}}\lambda_{pq}\ with\ r = \frac{b}{a}$ a. If $a = b$ or $r = 1$, then the invarianceness holds for every $\tilde{\lambda}_{pq} = \lambda_{pq}$ b. Otherwise, if $a \neq b$, then the expression for scaling becomes: $\tilde{\eta}_{pq} = \frac{\tilde{\mu}_{pq}(\tilde{\mu}_{pq})^{i+1}}{\tilde{\mu}_{p+1,0}\tilde{\mu}_{0,q+1,}}, with\ \acute{\mu}_{pq} = \frac{1}{a^{p+1}b^{q+1}}\mu_{pq}$
Feng and Keane (1994)	The proposed technique is called Aspect Moment Invariant (AMI). The technique eliminates the need of size normalization. The dynamic range remains constant with moment order. Mathematical expression: $\eta_{pq} = \frac{\mu_{00}^{\frac{p+q}{2}+1}}{\mu_{20}^{\frac{p+q}{2}}\mu_{02}^{\frac{p+q}{2}}}\mu_{pq}$
Raveendran, *et al.* (1995a,b)	The mathematical expression is different from that of (Raveendran, P., S. Jegannathan, et al., 1993). This equation is not rotation invariant. Mathematical expression: $\gamma_{pq} = \frac{\eta_{pq}}{\eta_{p+1,q+1}}, with\ \eta_{pq} = \left(\frac{b}{a}\right)^{\frac{q-p}{2}}$
Palaniappan, *et al.* (2000)	The expression for scale factor uses higher order but a second order moment equation is chosen. The expression is invariant

	to translation and non-uniform scaling. Mathematical expression: $$\eta_{pq} = \left(\frac{\tilde{\mu}_{02}}{\tilde{\mu}_{20}}\right)^{\frac{p-q}{4}} \tilde{\eta}_{pq}$$
Shamsuddin, *et al*. (2000)	An integration of Higher Order Moment and Aspect Moment Invariant. Mathematical expression: $$\eta_{pq} = \left(\frac{\mu_{20}{}^{(p+1)/2}\mu_{02}{}^{(q+1)/2}}{\mu_{40}{}^{(p+1)/2}\mu_{04}{}^{(q+1)/2}}\right)\mu_{pq}$$
Yinan, *et al.* (2003)	The expression can cover invariant for both region and boundary in discrete and continuous region. Mathematical expression: $$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^{\frac{p+q+2}{2}}}$$ $$\eta'_{pq} = \rho^{p+q}\eta_{pq} = \frac{\rho^{p+q}}{(\mu_{00})^{\frac{p+q+2}{2}}}\mu_{pq}$$ $$\eta"_{pq} = \frac{\mu_{pq}}{(\mu_{00})^{\frac{p+q+1}{2}}}$$
Muda, *et al.* (2007)	An integration of United Moment Invariant and Higher Order Centralized Scale Invariant.
Muda, *et al.* (2008)	An integration of Aspect Moment Invariant and United Moment Invariant.
Pamungkas and Shamsuddin (2009)	An integration of Aspect Moment Invariant and Weighted Central Moment.
Hameed, *et al.* (2014)	Non-Zero Invariants

Figure 1.1 illustrates the scenarios leading to the research problem in human action recognition. The action classification is crucial in recognizing the human action because the extracted features of the motion must be classified accordingly.

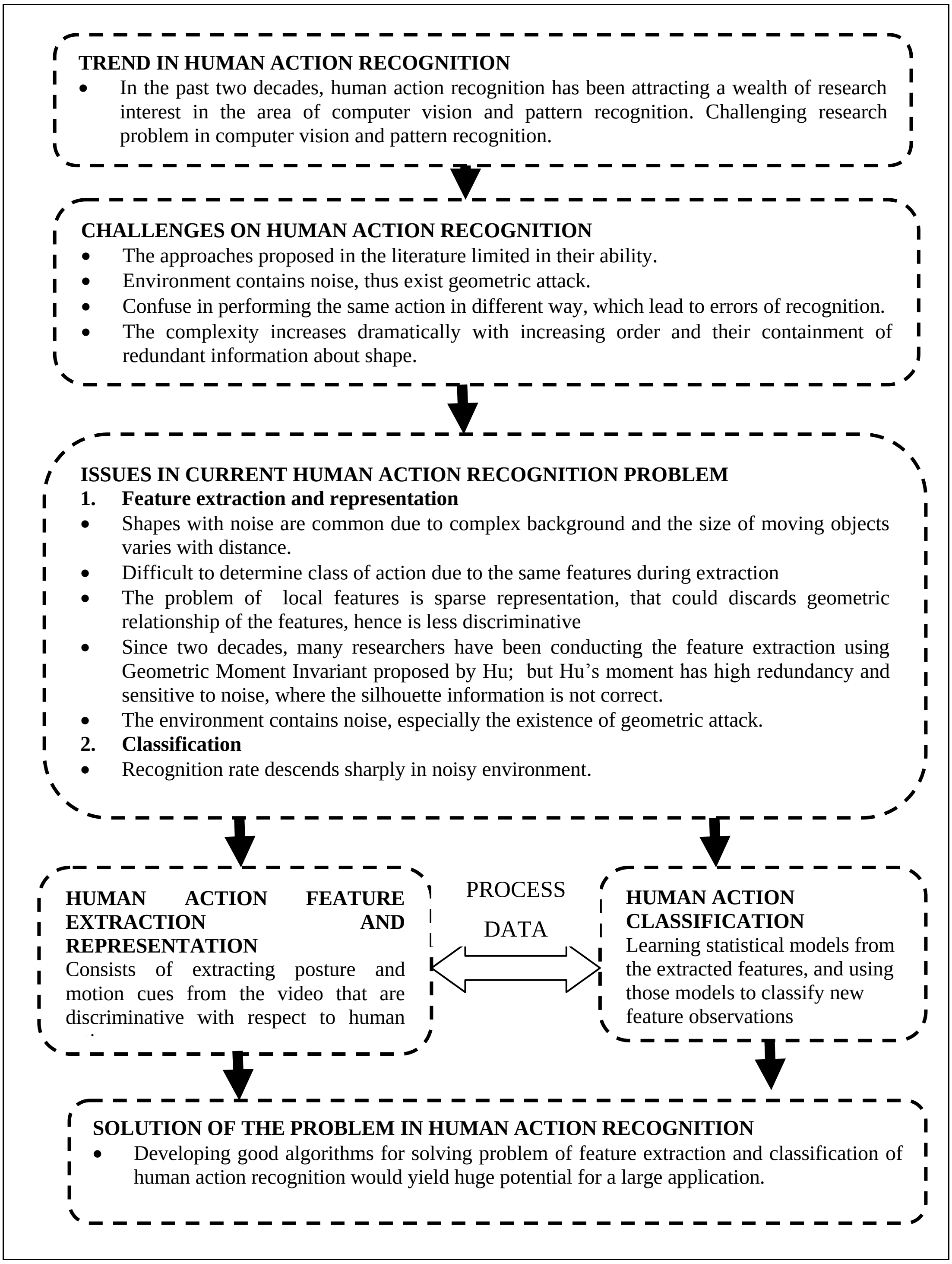

Figure 1.1 Scenarios leading to the Research Problem

In order to interpret the human behavior, it is important to recognize the human action. However, the interpretation still remains a challenging and complex task due to illumination and scale/view changes, occlusions and background clutter in the videos (Shabani, *et al.*, 2013; Wu, *et al.*,2011). Random movements cause scattered data, and consequently, these problem cause the features are not labeling action class correctly and representation of data features are not standardize properly (Mokhber, *et al*., 2008; Sen Wang, *et al.*, 2012; Shabani, *et al.*, 2013). The most important characteristic of invariance of the human action is the stability under different views but distinctive from the different classes. However, the increasing number of action classification in a video will lead to more challenges due to the higher overlapping between classes. This shows that inappropriate feature extraction and representation may directly cause a low accuracy in classification of human action. Therefore, more standardize uniform representation of data distributions are needed for recognition of human action.

1.3 Statement of the Problem

Some issues have been identified with regards to feature extraction and recognition task of human action in a video surveillance. In this study, there are two problem statements and potential solutions have been addressed as following:

Related issue 1: There is a large variation in the performance of human action, such as variation in human motion silhouette. Geometric Moment Invariant proposed by Hu's moment has high redundancy and sensitive to noise, where the silhouette information is not correct. Due to the environment contains noise; the features could be existence of geometric attack.

Hypothesis 1: The unique global features based on moment invariant algorithm are implemented in the process of feature extraction in order to improve the invarianceness of the human action feature vector representation. A good feature extraction approach should be able to generalize over variations within class (intra-class) and distinguish between actions of different classes (inter-class). Therefore, the variation of the features of human action can be minimize variation for intra-class and maximize variations for inter-class for human action.

Related issue 2: The orientation of a person might be different (such as front view or side view), and the motion direction might be changed when the person moves. Therefore, objects would result in a poor human body representation because the rate at which the actions are recorded affects the temporal extent of action, especially when motion features are used.

Hypothesis 2: A suitable human action representation should be utilized when dealing with these changes without affecting the recognition accuracy of action. Thus, the standard representation procedure for the effective learning methods is applied in order to raise the performance of the system and reduce the amount of accessible information to a manageable size of feature vector representation without losing any valuable information.

The research questions that need to be addressed in order to complement statement above are:

i. How to extract the robust image feature vector from geometrical moment function?
ii. How to determine the invarianceness of human motion from the geometrical feature vectors?
iii. How to standardize the uniqueness of the human action for effective learning methods?

1.4 Purpose of the Study

The purpose of the study is to develop a new scheme of human action recognition representation techniques based on the global features that can be achieved by extracting the silhouette of the human global features using the proposed geometrical moment invariant technique. The technique for enhancing the learning process is proposed to improve the variance between features that could increase the performance accuracy in human action recognition. This can be done by extracting the human silhouette using the proposed geometrical based moment, which include Higher Order Scale Moment Invariant (HGMI), Aspect Moment Invariant (AMI), United Moment Invariant (UMI), Aspect United Moment Invariant (AUMI), and Higher United Moment Invariant (HUMI). The data features will then be further validated using the invariant discretization for classification of human action recognition.

1.5 Objectives of the Study

The purpose of the study can be achieved with the following objectives:

i. To propose a new scheme for the human action recognition.
ii. To propose feature extraction based on geometrical moment invariant method for human action invarianceness.
iii. To propose compactness standard representation of geometrical features through supervised invariant discretization.

1.6 Scope of the Study

The scope of the study is limited to the following conditions:

i. The scope of feature extraction is only for two dimensional (2D) images silhouette.
ii. Sample image frames are obtained from IXMAS dataset. The dataset contain 13 actions (check watch; cross arms; scratch head; sit down; get up; turn around; walk; wave, punch; kick; point; pick up, throw).
iii. The size of resolution of each video image frame is 200 x 200 pixels.
iv. This research is focused on geometrical moment approaches to handle the feature representation of the human action obtained from video images.

v. This study will explore the substantially on feature extraction and representation; and feature interval for learning phase ONLY, and classification algorithm used for recognition will not be considered.

vi. Microsoft Studio 2010 is used to develop the proposed method.

vii. In this study, Microsoft excel, WEKA 3.7 is used to recognize the proposed method.

viii. SPSS toolkit and Minitab 15 is used as the statistical analysis to further validate the action recognition result.

ix. ANOVA is used for significant test for the proposed methods.

1.7 Research Methodology

This section briefly introduces the research methodology implemented in this study. The detailed discussion of the research methodology is presented in Chapter 3. Two phases are involved in order to achieve the goal and objectives of the study. Phase I is the investigation phase, in which the problem definition and formulation are explained, collection of data and the related literatures and existing methods are reviewed. Phase II is the development phase, which consists of two main stage: feature extraction and representation stage, standardize action learning features for classification stage, as shown in Figure 1.2. Five algorithm of the moment invariant for feature representation techniques are developed in this phase. The algorithm includes higher order scale Moment Invariant (HGMI), Aspect Moment Invariant (AMI), United Moment Invariant (UMI), Aspect United Moment Invariant (AUMI) and Higher United Moment Invariant (HUMI). The idea behind the proposed geometrical moment techniques is to extract the global features from human silhouette. After feature extraction, the human action invarianceness method is analyzed based on intra-class and inter-class variance, following which the extracted features vectors are discretized for performance in action learning. The invariant discretization is proposed in this action learning process in order to represent and illustrate the human action features in systematic representations. The final step is the classification of the human action.

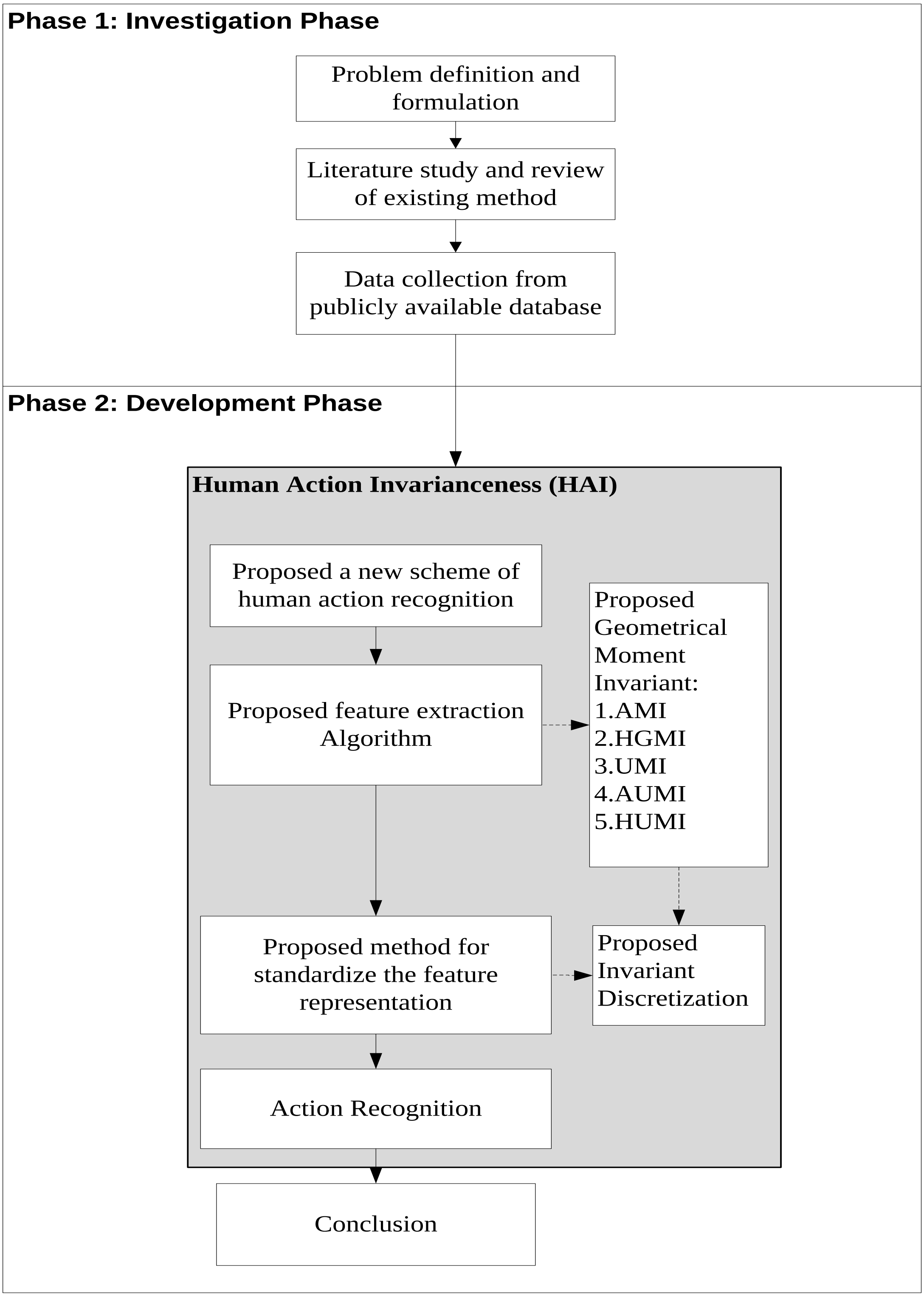

Figure 1.2 Overview of the research design

1.8 Research Contribution

This section highlights the research contribution that leads to the philosophy of the study in problem domain perspective. In this study, the focus is on the feature representation of human action using human silhouette as the primitive descriptive unit. The extracted global features from the proposed Geometrical Moment representation consist of human action features that can be generalised as the action features. For feature representation perspective, five types of geometrical moment are proposed as a technique to generate the representation for every video image in order to reduce the possible error. Then, for action learning perspective, these features are discretized using invariant discretization. The summary of the philosophy of the study is illustrated in Figure 1.4.

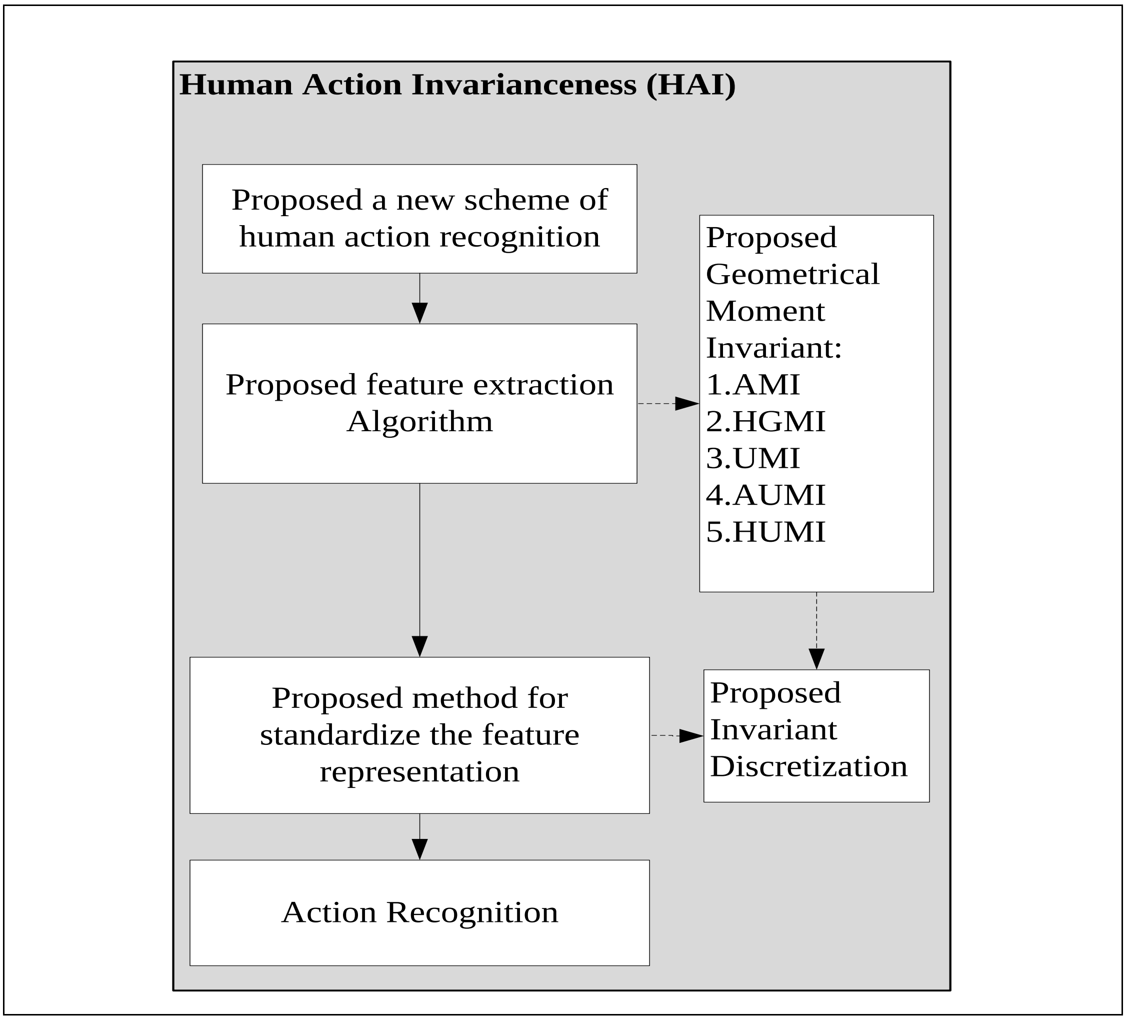

Figure 1.3 The philosophy of study

1.9 Organization of the Thesis

This thesis is organized into seven chapters. Chapter 1 describes the overview of chapter, problem background and problem statement, purpose and objectives of the study, research scope, methodology and contribution.

Chapter 2 consists of literature reviews in which previous works related to the issues were surveyed. This chapter consists of the concept of video surveillance, the human motion analysis model, research issues in human action recognition, feature extraction for human action recognition, and the foundations of moment invariant for feature extraction that includes the concepts and types of moment invariant. This is followed with the explanation on the discretization in action learning and the previous studies related to action learning. The explanation on the classifier that is used in the study is also included in the chapter, and finally the previous research on moment invariant in human action recognition is discussed.

The methodology of the research is explained in Chapter 3 that covers the overall research methodologies, which include the investigation phase and development phase. This chapter also provides the operational procedures used, covering data source collection procedure (database, feature extraction technique, data representation, and discretization procedure), performance measurement procedure (intra-inter class definition and ANOVA), and its development tools.

In Chapter 4 and Chapter 5 are the discussions of the study objectives as has been outlined in Chapter 1. Each chapter discusses the overall research findings for each objective. Lastly, Chapter 6 presents the conclusion of the study and recommendations for future research.

REFERENCES

Abas, K. H., Ono, O., and Khairul Hamimah A., 2010. Implementation of Multi Centroid Moment Invariants in Thermal-Based Face Identification System. *American Journal of Applied Sciences,Science Publications.* 7(3), 283-289.

Abas, K. H., and Osamu, O., 2009. Infrared-based face identification system via Multi-Thermal moment invariants distribution. *3rd International Conference on Signals, Circuits and Systems (SCS).*1-5.

Abdul, M. S. R., Basima, Z., Yacob., and Danny, T., Baito., 2013. Moment Invariants Based Feature Extraction For Classification of Syriac Alphabet Language. *International Journal of Advances in Engineering & Technology (IJAET),* 6(3), 1442-1451.

Abu Bakar, N., 2011. Integrated Zernike Formulation for Features Granularity in Pattern Recognition. *Master Thesis. Universiti Teknologi Malaysia, Johor Bahru.*

Abu Bakar, N., Shamsuddin, S. M., and Ali, A., 2010. An Integrated Formulation of Zernike Representation in Character Images. *Trends in Applied Intelligent Systems,* 6098, 359-368.

Abu Bakar, N. and Shamsuddin, S. M., 2009. United Zernike Invariants for Character Images.*Visual Informatics: Bridging Research and Practice*. 5857, 498-509.

Achard, C., Qu, X., Mokhber, A., and Milgram, M., 2008. A novel approach for recognition of human actions with semi-global features. *Machine Vision and Applications.* 19(1), 27-34.

Aggarwal, J. K., and Cai, Q.,1999. Human Motion Analysis: A Review. *Computer Vision and Image Understanding.* 73(3), 428-440.

Aggarwal, J. K., Cai, Q., Liao, W., and Sabata, B.,1994. Articulated and elastic non-rigid motion: a review. *Proceedings of the 1994 IEEE Workshop on the Motion of Non-Rigid and Articulated Objects.* 2-14.

Aggarwal, J. K., Cai, Q., Liao, W., and Sabata, B., 1998. Nonrigid Motion Analysis: Articulated and Elastic Motion. *Computer Vision and Image Understanding*. 70(2), 142-156.

Aggarwal, J. K., and Sangho, P., 2004. Human motion: modeling and recognition of actions and interactions. *Proceedings 2nd International Symposium on the 3D Data Processing, Visualization and Transmission* (*3DPVT).* 640-647

Agre, G., and Peev, S., 2002. On Supervised and Unsupervised Discretization *Cybernatixs And Information Technologies*. 2(2), 43-57.

Agresti, A., 2007. Chapter 12 Comparing Groups: Analysis of Variance (ANOVA) Methods from http://www.stat.ufl.edu/~aa/sta6127/ch12.pdf.

Ahad, M., Tan, J. K., Kim, H. S., and Ishikawa, S., 2008. Human activity recognition: Various paradigms. *International Conference on the Control, Automation and Systems* (*ICCAS).*1896-1901.

Ahad, M. A. R., Ogata, T., Tan, J.K., Kim, H., and Ishikawa, S., 2007. A smart automated complex motion recognition technique. *ACCV Workshop on Multi-dimensional and Multi-viewImage Processing*. 142-149.

Ahad, M. A. R., T. Ogata J.K. Tan, H.S. Kim, and Ishikawa, S., 2008a. Moment-based Human Motion Recognition from the Representation of DMHI Templates. *SICE Annual Conference*. 578-583.

Ahad, M. A. R., T. Ogata J.K. Tan, H.S. Kim, and Ishikawa, S., 2008b. View-based Human Motion Recognition in the Presence of Outliers. *Biomedical fuzzy and human sciences : the official journal of the Biomedical Fuzzy Systems Association.*13(1), 71-78.

Ahad, M. A. R., Tan, J., Kim, H., and Ishikawa, S., 2010. Action recognition by employing combined directional motion history and energy images. *IEEE Computer Society Conference on the Computer Vision and Pattern Recognition Workshops (CVPRW).* 73-78.

Ahad, M. A. R., Tan, J., Kim, H., and Ishikawa, S., 2011. Action dataset : A survey. *Proceedings of the SICE Annual Conference (SICE)*. 1650-1655.

Ahad, M. A. R., Tan, J. K., Kim, H., and Ishikawa, S., 2012. Motion history image: its variants and applications. *Machine Vision and Applications*. 23(2), 255-281.

Ahad, M. A. R., Tan, J. K., Kim, H. S., and Ishikawa, S., 2009. Temporal motion recognition and segmentation approach. *International Journal of Imaging Systems and Technology*. 19(2),91-99.

Ahad, M. D. A. R., Tan, J., Kim, H., and Ishikawa, S., 2010. Lower-Dimensional Feature Sets for Template-Based Motion Recognition Approaches. *Journal of Computer Science*. 6(8), 920-927.

Ahmad, M., and Lee, S.-W., 2008a. Human action recognition using shape and CLG-motion flow from multi-view image sequences. *Pattern Recognition*.41(7), 2237-2252.

Ahmad, M., and Lee, S.-W., 2008b. Recognizing human actions based on silhouette energy image and global motion description. *8th International Conference on Automatic Face and Gesture Recognition*. 1-6.

Ahmad, I., He, Z., Liao, M., Pereira, F., and Sun, M. T., 2008. Special Issue on Video Surveillance . *IEEE Transactions on Circuits and Systems for Video Technology*.18(8), 1001-1005.

Ahmad, M., and Hossain, M. Z., 2008. SEI and SHI representations for human movement recognition. *11th International Conference on on Computer and Information Technology (ICCIT)*.521-526.

Ali, S. S., Zafar, M. F., and Tayyab, M., 2009. Detection and Recognition of Human in Videos Using Adaptive Method and Neural Net. *International Conference on the Soft Computing and Pattern Recognition (SOCPAR)*. 604-609.

Anh-Phuong, T., Wolf, C., Lavoue, G., Baskurt, A., and Jolion, J., 2010. Pairwise Features for Human Action Recognition. *Proceedings of the 2010 20th International Conference on Pattern Recognition*. 3224-3227.

Antonakaki, P., Kosmopoulos, D., and Perantonis, S. J., 2009. Detecting abnormal human behaviour using multiple cameras. *Signal Processing*. 89(9), 1723-1738.

Anwaar ul, H., Gondal, I., and Murshed, M., 2011. On dynamic scene geometry for view-invariant action matching. *IEEE Conference on the Computer Vision and Pattern Recognition (CVPR)*. 3305-3312.

Arantes, M., and Gonzaga, A., 2011. Human gait recognition using extraction and fusion of global motion features. *Multimedia Tools and Applications,* 55(3), 655-675.

Balslev, I., Døring, K., and Eriksen, R. D., 2000. Weighted central moments in pattern recognition. *Pattern Recognition Letters*. 21(5), 381-384.

Barnachon, M., Bouakaz, S., Boufama, B., and Guillou, E., 2014. Ongoing human action recognition with motion capture. *Pattern Recognition*. 47(1), 238-247.

Belkasim, S. O., Shridhar, M., and Ahmadi, M., 1989. Shape Recognition Using Zernike Moment Invariants. *23rd Asilomar Conference on Signals, Systems and Computers*. l(1), 167-171.

Bing, S., Junchi, Y., and Yuncai, L., 2010. Human Gait Recognition by Integrating Motion Feature and Shape Feature. *International Conference on the Multimedia Technology (ICMT).* 1-4.

Blank, M., Gorelick, L., Shechtman, E., Irani, M., and Basri, R., 2005. Actions as space-time shapes. *10th IEEE International Conference on Computer Vision (ICCV).* l(2), 1395-1402.

Bloom, V., Argyriou, V., and Makris, D., 2013. Dynamic Feature Selection for Online Action Recognition. *Human Behavior Understanding*. 8212, 64-76)

Bobick, A., 1997. Movement, activity and action: the role of knowledge in the perception of motion. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*. 352(1358), 1257-1265.

Bobick, A., and Davis, J., 1996. An appearance-based representation of action. *Proceedings of the 13th International Conference on Pattern Recognition.* 1, 307-312.

Bobick, A. F., and Davis, J. W., 2001. The recognition of human movement using temporal templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 23(3), 257-267.

Botao Wong, Jiang Sun, and Cai, A., 2002. Relative Moments and their Application to Geometric Shape Recognition. *Journal of Image and Graphics*. 6(3), 296-300.

Bouziane, A., Chahir, Y., Molina, M., and Jouen, F., 2013. Unified framework for human behaviour recognition: An approach using 3D Zernike moments. *Neurocomputing*. 100, 107-116.

Bradski, G., and Davis, J., 2002. Motion segmentation and pose recognition with motion history gradients. *Machine Vision Application*. 13(3), 174–184.

Bregler, C., 1997. Learning and recognizing human dynamics in video sequences. *Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 568-574.

Brown, S. D., and Myles, A. J., 2013. Decision Tree Modeling in Classification. *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering.* 3, 541-546.

Buxton, H., and Gong, S., 1995. Visual surveillance in a dynamic and uncertain world. *Artificial Intelligence*. 78(1-2), 431-459.

Wren, C., Sparacino, F., Azarbayejani, A., T. Darrell, T. S., Chao, K. A. C., Hlavac, M., Russell, K., and Pentland, A., 1997. Perceptive Spaces for Peformance and Entertainment: Untethered Interaction Using Computer Vision and Audition. *Applied Artificial Intelligence*. 11(4), 267–284.

Calderara, S., Cucchiara, R., and Prati, A., 2008. Action Signature: A Novel Holistic Representation for Action Recognition. *IEEE Fifth International Conference on the Advanced Video and Signal Based Surveillance* (*AVSS '08)*. 121-128.

Cédras, C., and Shah, M.., 1995. Motion-based recognition a survey. *Image and Vision Computing.* 13(2), 129-155.

Chaaraoui, A. A., Climent-Pérez, P., and Flórez-Revuelta, F., 2013. Silhouette-based human action recognition using sequences of key poses. *Pattern Recognition Letters.* 34(15), 1799-1807.

Chaquet, J. M., Carmona, E. J., and Fernández-Caballero, A., 2013. A survey of video datasets for human action and activity recognition. *Computer Vision and Image Understanding.* 117(6), 633-659.

Chen, C.-C., 1993. Improved moment invariants for shape discrimination. *Pattern Recognition*. 26(5), 683-686.

Chen, R., Wang, S., Gong, L., and Liu, C., 2013. Hand Gesture Recognition for Human-Computer Interaction Using Moment Invariants and Neural Network Classifier. *Intelligent Robotics and Applications*. 8102, 661-667.

Cheng, J., Liu, H., and Li, H, 2014. Silhouette analysis for human action recognition based on maximum spatio-temporal dissimilarity embedding. *Machine Vision and Applications*. 25(4), 1007-1018.

Chenglong, Y., Xuan, W., Hejiao, H., Jianping, S., and Kun, W., 2010. Vision-Based Hand Gesture Recognition Using Combinational Features. *Sixth*

International Conference on the Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP). 543-546.

Cherla, S., Kulkarni, K., Kale, A., and Ramasubramanian, V., 2008. Towards fast, view-invariant human action recognition. *IEEE Computer Society Conference on the Computer Vision and Pattern Recognition Workshops (CVPRW '08)*. 1-8.

Clawson, K., Jing, M., Scotney, B., Wang, H., and Liu, J., 2014. Human Action Recognition in Video via Fused Optical Flow and Moment Features – Towards a Hierarchical Approach to Complex Scenario Recognition. *MultiMedia Modeling*. 8326, 104-115.

Concha, O. P., Xu, R. Y. D., and Piccardi, M., 2010. Robust Dimensionality Reduction for Human Action Recognition. *International Conference on the Digital Image Computing: Techniques and Applications (DICTA).* 349-356.

Cook, D., Feuz, K., and Krishnan, N., 2013. Transfer learning for activity recognition: a survey. *Knowledge and Information Systems.* 36(3), 537-556.

Costa, J. A. F., and De Andrade Netto, M. L., 1997. Parts classification in assembly lines using multilayer feedforward neural networks. *Computational Cybernetics and Simulation IEEE International Conference on the Systems Man, and Cybernetics*. 4, 3872-3877.

D. Marr, and Vaina, L., 1982. Representation and recognition of the movements of shapes. *Philosophical Transactions of the Royal Society of London B 214*. 501-524.

Dai, S., Huang, H., Zhangying, G., Kai, L., and Shumei, X., 2009. Vehicle-logo Recognition Method Based on Tchebichef Moment Invariants and SVM, *WRI World Congress on the Software Engineering (WCSE '09)*. 18-21.

Darrell, T., Blumberg, B., Daniel, S., Rhodes, B., Maes, P., and Pentland, A., 1995. Alive: Dreams and Illusions. *ACM SIGGraph, Computer Graphics Visual Processing*.

Davis, J., and Bradski, G., 1999. Real-time motion template gradients using Intel CVLib. *Proceeding ICCV Workshop on Frame-Rate Vision*. 1-20.

Deghan, M., and Faez, K., 1997. Farsi Handwritten Character Recognition with Moment Invariants. *13th International Conference on Digital Signal Processing Proceeding (DSP)*. 97(2), 507-510.

Di, W., and Ling, S., 2013. Silhouette Analysis-Based Action Recognition Via Exploiting Human Poses. *IEEE Transactions on Circuits and Systems for Video Technology*. 23(2), 236-243.

Diao, Z., Zheng, A., and Wu, Y., 2012. Shape Feature Extraction of Wheat Leaf Disease Based on Invariant Moment Theory. *Computer and Computing Technologies in Agriculture V*. 369, 168-173.

Ding, M., Chang, J., and Peng, J., 1992. Research on Moment Invariant Algorithm", *Journal of Data Acquisition & Processing*. 7(1), 1-9.

Donderac, R., Doermann, D., and Davis, L., 2009. Action recognition based on human movement characteristics. *Proceedings of the International conference on Motion and video computing.*1-8.

Dunlap, B. I. and Shashi, and Karna, P, 2006. Dipole Moments from Atomic-Number-Dependent Potentials in Analytic Density-Functional Theory. *Journal of Chemical Physics*. 125, 1-9.

Everts, I., van Gemert, J. C., and Gevers, T., 2014. Evaluation of Color STIPs for Human Action Recognition. *IEEE Transactions on Image Processing*.23(4), 1569-1580.

Farokhi, S., Shamsuddin, S. M., Sheikh, U. U., Flusser, J., Khansari, M., and Jafari-Khouzani, K., 2014. Near infrared face recognition by combining Zernike moments and undecimated discrete wavelet transform. *Digital Signal Processing*. 31, 13-27.

Fayolle, J., Ducottet, C., Fournel, T., and Schon, J. P., 1996. Motion characterization of unrigid objects by detecting and tracking feature points. *Proceedings International Conference on Image Processing*. 3,803-806.

Feng, P., and Keane, M., 1994. A new set of moment invariants for handwritten numeral recognition. *Proceedings IEEE International Conference Image Processing (ICIP-94)*. 1, 154-158.

Flusser, J., and Suk, T., 1993. Pattern recognition by affine moment invariants", *Pattern Recognition*.26(1), 167-174.

Flusser, J., Suk, T., and Zitová, B., 2009. Moments and Moment Invariants in Pattern Recognition: Wiley & Sons Ltd.

Gama, J., 2001. Functional trees for classification. *Proceedings IEEE International Conference on Data Mining (ICDM 2001)*, 147-154.

Gao, H., Liu, X., Liu, Z., and Hong, K., 2012. Investigation on Traffic Signs Recognition Based on BP Neural Network and Invariant Moments. *Advances in Future Computer and Control Systems.* 159, 157-162.

Gavrila, D. M., 1999. The Visual Analysis of Human Movement: A Survey. *Computer Vision and Image Understanding.* 73(1), 82-98.

Ghodke, S., and Baldwin, T., 2007. An Investigation into the Interaction Between Feature Selection and Discretization: Learning How and When to Read Numbers. *Advances in Artificial Intelligence*. 4830, pp. 48-57.

Gilbert, A., Illingworth, J., and Bowden, R., 2008. Scale Invariant Action Recognition Using Compound Features Mined from Dense Spatio-temporal Corners. *Proceedings of the 10th European Conference on Computer Vision: Part I.* 222-233.

Gorelick, L., Galun, M., Sharon, E., Basri, R., and Brandt, A., 2006. Shape Representation and Classification Using the Poisson Equation. *Transactions on Pattern Analysis and Machine Intelligence.* 28(12), 1991-2005.

Gurrapu, C., 2004. Human Action Recognition In Video Data For Surveillance Applications. Masters by Research Thesis. Queensland University of Technology.

H. Qu, L. Wang, and Leckie, C., 2010. Action recognition using space-time shape difference images. *20th International Conference on Pattern Recognition (ICPR)*. 3661-3664.

Hameed, V. A., Shamsuddin, S. M., Darus, M., and Ralescu, A. L., 2014. Faster Computation of Non-Zero Invariants from Graph Based Method. *Neurocomputing*. 137, 96-104.

Haritaoglu, I., Harwood, D., and David, L. S., 2000. W4: Real-Time Surveillance of People and Their Activities. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 22(8), 809-830.

Haritaoglu, I., Harwood, D., and Davis, L. S.,1998. Who? When? Where? What? A real time system for detecting and tracking people. *Proceedings 3rd IEEE International Conference on Automatic Face and Gesture Recognition*. 1-6.

Hongtao, L., and Jicheng, Z., 1993. Research on Discrete Moment Invariant Algorithm. *Journal of Data Acquisition and Processing.* 8(2), 151-155.

Hota, R. N., Venkoparao, V., and Rajagopal, A., 2007. Shape Based Object Classification for Automated Video Surveillance with Feature Selection.

10th International Conference on Information Technology, (ICIT 2007). 97-99.

Hu, M.-K. (1962). Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*. 8(2), 179-187.

Huan, L., Farhad, H., Chew, L., Tan., , and Manoranjan, D., 2002. Discretization: an enabling technique. *Data Minning Knowledge Discovery*. 6, 393–423.

Ikizler, N., Cinbis, R. G., and Duygulu, P., 2008. Human action recognition with line and flow histograms. *19th International Conference on Pattern Recognition (ICPR 2008)*. 1-4.

Iosifidis, A., Tefas, A., Nikolaidis, N., and Pitas, I., 2012. Multi-view human movement recognition based on fuzzy distances and linear discriminant analysis. *Computer Vision and Image Understanding.* 116(3), 347-360.

Iwashita, Y., and Kurazume, R., 2009. Person identification from human walking sequences using affine moment invariants. *IEEE International Conference on Robotics and Automation (ICRA '09)*. 436-441.

Jin Soo, N., and Kang Hyeon, R., 2005. Palmprint identification algorithm using Hu invariant moments and Otsu binarization. *4th Annual ACIS International Conference on Computer and Information Science*. 94-99.

Jingen, L., Shah, M., Kuipers, B., and Savarese, S., 2011. Cross-view action recognition via view knowledge transfer. *IEEE Conference on the Computer Vision and Pattern Recognition (CVPR)*. 3209-3216.

Junejo, I., Junejo, K., and Aghbari, Z., 2013. Silhouette-based human action recognition using SAX-Shapes. *The Visual Computer*. 30(3), 259-269.

Kaiqi, H., Yeying, Z., and Tieniu, T., 2012. A Discriminative Model of Motion and Cross Ratio for View-Invariant Action Recognition. *IEEE Transactions on Image Processing*. 21(4), 2187-2197.

Kale, K. V., Deshmukh, P. D., Chavan, S. V., Kazi, M. M., and Rode, Y. S., 2013. Zernike moment feature extraction for handwritten Devanagari compound character recognition. *Science and Information Conference (SAI)*. 459-466

Kanade, T., Collins, R. T., Lipton, A. J., Burt, P. J., and Wixson, L., 1998. Advances in Cooperative Multi-Sensor Video Surveillance. *Proceedings of DARPA Image Understanding Workshop*.3-24.

Kanade, T., Collins, R., Lipton, A., Anandan, P., and Burt, P., 1997. Cooperative multisensor video surveillance. *In Proceedings of the 1997 DARPA Image Understanding Workshop.* 1, 3–10.

Ke, S.-R. T., H.L.U.; Lee, Y.-J.; Hwang, J.-N.; Yoo, J.-H.; and Choi, K.-H., 2013. A Review on Video-Based Human Activity Recognition. . *Computers.* 2(88-131).

Kim, W. Y., and Po,Y., 1994. A Practical Pattern Recognition System for Translation, Scale And Rotation Invariance. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Proceedings (CVPR'94).* 391-396.

Ko, T., 2008. A survey on behavior analysis in video surveillance for homeland security applications. *37th IEEE Applied Imagery Pattern Recognition Workshop (AIPR'08)*. 1-8.

Koker, R., Oz, C., and Ferikoglu, A., 2001. Development of a vision based object classification system for an industrial robotic manipulator. *The 8th IEEE International Conference on.the Electronics, Circuits and Systems (ICECS).* 3, 1281-1284.

Kui, J., and Dit-Yan, Y., 2008. Human action recognition using Local Spatio-Temporal Discriminant Embedding. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2008)*. 1-8.

Kulkarni, A. D., Yap, A. C., and Byars, P., 1990. Neural networks for invariant object recognition. Proceedings of the 1990 Symposium on Applied Computing. 28-32.

Kulkarni, S., Borde, P., Manza, R., and Yannawar, P., 2014. Recognition of Handwritten MODI Numerals using Hu and Zernike features. arXiv preprint arXiv:1404.1151. 1-7.

Kumar, S., Kumar, D. K., Sharma, A., and McLachlan, N., 2004. Classification of hand movements using motion templates and geometrical based moments. *Proceedings of International Conference on.Intelligent Sensing and Information Processing*. 299-304

L. Wang, and Leckie, C., 2010. Encoding actions via the quantized vocabulary of averaged silhouettes. *20th International Conference Pattern Recognition*. 3657-3660.

Lara, O. D., and Labrador, M. A., 2013. A Survey on Human Activity Recognition using Wearable Sensors. *Communications Surveys & Tutorials, IEEE.* 15(3), 1192-1209.

Lassoued, I., Zagrouba, E., and Chahir, Y., 2011. An Efficient Approach for Video Action Classification Based on 3D Zernike Moments. *Future Information Technology*. 185, 196-205.

Leng, W. Y., Shamsuddin, S. M., and Srihari, S. N., 2013. Fusing Modalities in Forensic Identification with Score Discretization. AFHA. 46-50.

Li Ping Zhang, C. W., Hong Zhang., and Bo Zhang ., 2010. Aircraft Type Recognition in High-Resolution SAR Images Using Multi-Scale Auto Convolution. *Key Engineering Materials.* 439-440, 1475-1480.

Li, R. F., Wang, L. L., and Zhang, T. F., 2014. Efficient feature representation for human action recognition in DSSP. *4th International Conference on Advances in Materials and Manufacturing (ICAMMP)*. 889-890, 1057-064.

Liao, S. X., and Pawlak, M., 1996. On image analysis by moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 18(3), 254-266.

Lihong, L., Dongli, J., Xiangguo, C., and Lixin, S., 2006. A Fast Discrete Moment Invariant Algorithm and Its Application on Pattern Recognition. *The Sixth World Congress on Intelligent Control and Automation (WCICA)*. 2, 9773-9777.

Little, J., and Boyd, J., 1996. Recognizing People by Their Gait: The Shape of Motion. *Videre*. 1, 1-32.

Liu, H., Feris, R., and Sun, M.-T., 2011. Benchmarking Datasets for Human Activity Recognition. *Visual Analysis of Humans.* 411-427.

Lu, Y., Li, Y., Shen, Y., Ding, F., Wang, X., Hu, and J., Ding, S., 2012. A Human Action Recognition Method Based on Tchebichef Moment Invariants and Temporal Templates. *4th International Conference on Intelligent Human-Machine Systems and Cybernetics*. 2, 76-79.

Majid, M., Altuwaijri, and Bayoumi, M. ., (1994). Arabic Text Recognition Using Neural Networks. *IEEE International Symposium on Circuit and Systems (ISCAS'94).* 6,415-418.

Marr, D., and Nishihara, H. K., 1978. Representation and recognition of the spatial organization of three-dimensional shapes. *Proc. R. Soc. Lond. B Biol Sci.* 200(1140), 269-294.

McAuley, A., Coker, A., and Saruhan, K., 1991. Effect of noise in moment invariant neural network aircraft classification. *Proceedings of the IEEE 1991 National Aerospace and Electronics Conference (NAECON)*. 2, 743-749.

Megavannan, V., Agarwal, B., and Venkatesh Babu, R., 2012. Human action recognition using depth maps. *International Conference on Signal Processing and Communications (SPCOM)*. 1-5.

Moeslund, T. B., and Erik, G., 2001. A survey of computer vision-based human motion capture. *Computer Vision Image Understanding*.81(3), 231-268.

Mohammed, B.O., and Shamsuddin, S.M., 2011. Feature Discretization for Individuality Representation in Twins Handwritten Identification. *Journal of Computer Science*. 7(7), 1080-1087.

Mohammed, B.O., and Shamsuddin, S.M., 2012. Improvement in twins handwriting identification with invariants discretization. *EURASIP Journal on Advances in Signal Processing*. 2012(48), 1-12.

Mokhber, A., Achard, C., and Milgram, M., 2008. Recognition of human behavior by space-time silhouette characterization. *Pattern Recognition Letters*. 29(1), 81-89.

Mokhber, A., Achard, C., Qu, X., and Milgram, M., 2005. Action Recognition with Global Features. *Computer Vision in Human-Computer Interaction*. 3766, 110-119.

Moussa, M. M., Hamayed, E., Fayek, M. B., and El Nemr, H. A., 2013. An enhanced method for human action recognition. *Journal of Advanced Research*.

Muda, A. K., Shamsuddin, S. M., and Darus, M., 2007a. Embedded Scale United Moment Invariant for Identification of Handwriting Individuality. *Computational Science and Its Applications* . 4705, 385-396.

Muda, A. K., Shamsuddin, S. M., and Darus, M., 2007b. Invarianceness of Higher Order United Scale Invariants. *International Journal of Advanced in Computer Science and Engineering*. 1(2), 105-118.

Muda, A. K., Shamsuddin, S. M., and Darus, M., 2008a. Discretization of integrated moment invariants for Writer Identification. *Proceedings of the Fourth IASTED International Conference on Advances in Computer Science and Technology*. 372-377.

Muda, A. K., Shamsuddin, S. M., and Darus, M., 2008b. Invariants Discretization for Individuality Representation in Handwritten Authorship. *Computational Forensics*. 5158, 218-228.

Muda, A. K., Shamsuddin, S. M., and Abraham, A., 2010. Improvement of Authorship invarianceness for individuality representation in writer identification. *Neural Netw World*. 3(10), 371–387.

Mukundan, R., and Ramakrishnan, K. R., 1998. Moment functions in image analysis theory and applications. Singapore; New Jersey; London: World scientific.

Nabatchian, A., Abdel-Raheem, E., and Ahmadi, M., 2008. Human Face Recognition Using Different Moment Invariants: A Comparative Study. *The Proceedings of the 2008 Congress on Image and Signal Processing*. 3, 661-666.

Nabatchian, A., Makaremi, I., Abdel-Raheem, E., and Ahmadi, M., 2008. Pseudo-Zernike Moment Invariants for Recognition of Faces Using Different Classifiers in FERET Database. *Third International Conference on.the Convergence and Hybrid Information Technology*. 1, 933-936.

Natarajan, P., and Nevatia, R., 2008. View and scale invariant action recognition using multiview shape-flow models. IEEE Conference on Computer Vision and Pattern Recognition. 1-8.

Nor'aini, A. J., Raveendran, P., and Selvanathan, N., 2006. Human Face Recognition using Zernike moments and Nearest Neighbor classifier. 4th Student Conference on the Research and Development (SCOReD). 120-123.

Oz, C., and Ercal, F., 2003. Signature recognition and verification with ANN using moment invariant method. *Proceedings of the Artificial Neural Networks in Engineering Conference Smart Engineering System Design: Neural Networks, Fuzzy Logic, Evolutionary Programming, Complex Systems and Artificial Life*. 195-202.

Palaniappan, R., Raveendran, P., and Omatu, S., 2000. New Invariant Moments for Non-Uniformly Scaled Images. *Pattern Analysis & Applications*. 3(2), 78-87.

Pamungkas, R., and Shamsuddin, S.M., 2009. Weighted Aspect Moment Invariant in Pattern Recognition. *Computational Science and Its Applications – ICCSA 2009*. 5593, 806-818.

Pamungkas, R. P., and Shamsuddin, S. M., 2009. Weighted Central Moment for Pattern Recognition: Derivation, Analysis of Invarianceness, and Simulation Using Letter Characters. *Third Asia International Conference on Modelling & Simulation (AMS '09)*. 102-106.

Parameswaran, V., and Chellappa, R., 2002. Quasi-invariants for human action representation and recognition. *16th International Conference on Pattern Recognition*. 1, 307-310.

Paulraj, M. P., Yaacob, S., Azalan, M. S. Z., and Palaniappan, R., 2011. A phoneme based sign language recognition system using 2D moment invariant interleaving feature and Neural Network. IEEE Student Conference on Research and Development (SCOReD). 111-116.

Phiasai, T., Arunrungrusmi, S., and Chamnongthai, K., 2001. Face recognition system with PCA and moment invariant method. IEEE International Symposium on Circuits and Systems (ISCAS 2001). 2, 165-168.

Polana, R., and Nelson, R., 1994. Low level recognition of human motion (or how to get your man without finding his body parts). *Proceedings of the IEEE Workshop on Motion of Non-Rigid and Articulated Objects*. 77-82.

Poppe, R., 2007. Vision-based human motion analysis: An overview. *Computer Vision Image Understanding.* 108(1-2), 4-18.

Poppe, R., 2010. A survey on vision-based human action recognition. *Image and Vision Computing.* 28(6), 976-990.

Puteh, S., 2004. Feature Extraction of Trademark Images Using Geometric Moment Invariant and Zernike Moment-A Comparison. *Chiang Mai Journal of Science.* 3, 217-222.

Qi, Z.-Q., Jing, X.-J., You, S.-Q., and Sun, H.-B., 2013. Application of invariant moment method for HAPS ISAR imaging with micro parts. *Tongxin Xuebao/Journal on Communications.* 34(8), 70-75.

Ramteke, R. J., 2010. Invariant Moments Based Feature Extraction for Handwritten Devanagari Vowel Recognition. *International Journal of Computer Application.* 1, 1-5.

Rao, C., and Shah, M., 2001. View-invariant representation and learning of human action. *Proceedings IEEE Workshop on Detection and Recognition of Events in Video.* 50(2),203-226.

Raveendran, P., Jegannathan, S., and Omatu, S., 1993. New regular moment invariants to classify elongated and contracted images. Proceedings of 1993 International Joint Conference on Neural Networks. 3, 2089-2092.

Raveendran, P., Jegannathan, S., and Omatu, S., 1994. Classification of elongated and contracted images using new regular moments. *IEEE International Conference on Neural Networks, IEEE World Congress on Computational Intelligence*. 6, 4154-4158.

Raveendran, P., and Omatu, S., 1995. A new technique to derive features for shift and unequally scaled images. *Proceedings., IEEE International Conference on Neural Networks*. 4, 2077-2080.

Raveendran, P., Omatu, S., and Abu Bakar, W., 1995. Neuro-pattern classification of elongated and contracted images. *Information Sciences - Applications.* 3(3), 209-221.

Raveendran, P., Omatu, S., and Poh Sin, C., 1997. A new technique to derive invariant features for unequally scaled images. *IEEE International Conference on Systems, Man, and Cybernetics,. Computational Cybernetics and Simulation*. 4, 3158-3163.

Raviraja, S., Osman, S. S., and Kardman., 2008. A Novel Technique For Malaria Diagnosis Using Invariant Moments And By Image Compression. *4th Kuala Lumpur International Conference on Biomedical Engineering*. 21, 730-733.

Regazzoni, C., Ramesh, V., and Foresti, G., 2001. Special issue on video communications, processing, and understanding for third generation surveillance systems. *Proceedings of the IEEE.* 89(10), 1355-1539.

Reiss, T. H., 1991. The revised fundamental theorem of moment invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 13(8), 830-834.

Ren, K., and Qu, J., 2014. Identification of Shaft Centerline Orbit for Wind Power Units Based on Hopfield Neural Network Improved by Simulated Annealing. *Hindawi Publishing Corporation. Mathematical Problems in Engineering.* 2014, 6.

Rosales, R., 1998. Recognition of Human Action Using Moment-Based Features. 10th International Multimedia Modeling Conference. 1.

Sabaei, M., and Faez, K., 1997. Unsupervised Classification of Hand Written Farsi Numerals Using Evolution Strategy. *Proceedings of TENCON-97, Queensland, Australia*. 403-406.

Sadek, S., Al-Hamadi, A., Elmezain, M., Michaelis, B., and Sayed, U., 2010. Human activity recognition via temporal moment invariants. *Proceedings of the The 10th IEEE International Symposium on Signal Processing and Information Technology*. 79-84.

Sadek, S., Al-Hamadi, A., Michaelis, B., and Sayed, U., 2010. Human Activity Recognition: A Scheme Using Multiple Cues. *Advances in Visual Computing.* 6454, 574-583.

Sadek, S., Al-Hamadi, A., Michaelis, B., and Sayed, U., 2012. Human action recognition via affine moment invariants. *21st International Conference on Pattern Recognition (ICPR)*. 218-221.

Samy Sadek, Ayoub Al-Hamadi, Gerald Krell, and Michaelis, B., 2013. Affine-Invariant Feature Extraction for Activity Recognition. *Machine vision.* 2013.

Sanossian, H. Y. Y., 1996. *An Arabic character recognition system using neural network.* Proceedings of the 1996 *IEEE Signal Processing Society Workshop Neural Networks for Signal Processing [1996] VI*. 340-348.

Sarfaz, M., Nawaz, S.N., and Al-Khuraidly, A., 2003. Offline Arabic Text Recognition System. *Proceedings. 2003 International Conference on Geometric Modeling and Graphics*. 30-35.

Sen Wang, Yi Yang, Zhigang Ma, Xue Li, Chaoyi Pang, and Hauptmann., A. G., 2012. Action Recognition by Exploring Data Distribution and Feature Correlation. *IEEE Conference Computer Vision and Pattern Recognition (CVPR)*. 1370-1377.

Shabani, A. H., Zelek, J. S., and Clausi, D. A., 2013. Multiple scale-specific representations for improved human action recognition. *Pattern Recognition Letters*. 34(15), 1771-1779.

Shamsuddin, S.M., Sulaiman, N., and Darus, M., 2000a. Feature extraction with an improved scale-invariants for deformation digits. *International Journal of Computer Mathematics.* 76(1), 13-23.

Shamsuddin, S.M., Sulaiman, N., and Darus, M., 2000b. Improved Scale Invariant Moment for Deformation Digits. *International Journal of Computer Mathematics.* 74(4), 439-447.

Shamsuddin, S.M., Sulaiman, N., and Darus, M., 2002. Invarianceness of higher order centralised scaled-invariants undergo basic transformations . *International Journal of Computer Mathematics* .79, 39-48.

Shao, L., Ji, L., Liu, Y., and Zhang, J., 2012. Human action segmentation and recognition via motion and shape analysis. *Pattern Recognition Letters. 33*(4), 438-445.

Sharma, A., Kumar, D. K., Kumar, S., and McLachlan, N., 2004. Recognition of human actions using moment based features and artificial neural networks. Proceedings. 10th International.Multimedia Modelling Conference. 368.

Shen, J., Yang, W., and Liao, Q., 2011. Multiview human pose estimation with unconstrained motions. *Pattern Recognition Letters.* 32(15), 2025-2035.

Singh, C., Walia, E., and Mittal, N., 2011. Rotation invariant complex Zernike moments features and their applications to human face and character recognition. *Computer Vision, IET.* 5(5), 255-265.

Starner, T., and Pentland, A., 1995. Real-time American Sign Language recognition from video using hidden Markov models. Proceedings. *International Symposium on Computer Vision*. 265-270.

Suganthan, P. N., Teoh, E. K., and Mital, D. P., 1993. *M*ultilayer backpropagation network for flexible circuit recognition. *International Conference on Proceedings of the Industrial Electronics, Control, and Instrumentation*. 3, 1645-1650.

Tan, T. N., Sullivan, G. D., and Baker, K. D., 1998. Model-Based Localisation and Recognition of Road Vehicles. *International Journal of Computer Vision.* 27(1), 5-25.

Theodorakopoulos, I., Kastaniotis, D., Economou, G., and Fotopoulos, S., 2013. Pose-based human action recognition via sparse representation in dissimilarity space. *Journal of Visual Communication and Image Representation*. 25(1), 12-23.

Tsai, C.-J., Lee, C.-I., and Yang, W.-P., 2008. A discretization algorithm based on Class-Attribute Contingency Coefficient. *Information Sciences.* 178(3), 714-731.

Parameswaran, V., and Chellappa, R., 2005. Human action-recognition using mutual invariants. *Computer Vision and Image Understanding.* 98, 295–325.

Vishwakarma, S., and Agrawal, A., 2013. A survey on activity recognition and behavior understanding in video surveillance. *The Visual Computer.* 29(10), 983-1009.

Viswanathan, P., Krishna, P. V., and Hariharan, S., 2010. Multimodal Biometric Invariant Moment Fusion Authentication System. *Information Processing and Management*. 70, 136-143.

Wang, L., Hu, W., and Tan, T., 2003. Recent developments in human motion analysis. *Pattern Recognition.* 36(3), 585-601.

Wanqing, L., Zhengyou, Z., and Zicheng, L., 2010. Action recognition based on a bag of 3D points. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 1-6.

Weiming, H., Tieniu, T., Liang, W., and Maybank, S., 2004. A survey on visual surveillance of object motion and behaviors. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews.* 34(3), 334-352.

Weinland, D., Boyer, E., and Ronfard, R., 2007. Action Recognition from Arbitrary Views using 3D Exemplars. IEEE 11th International Conference on Computer Vision. 1-7.

Weinland, D., Ronfard, R., and Boyer, E., 2006. Free viewpoint action recognition using motion history volumes. *Computer Vision Image Understanding.* 104(2), 249-257.

Weinland, D., Ronfard, R., and Boyer, E., 2011. A survey of vision-based methods for action representation, segmentation and recognition. *Computer Vision and Image Understanding.* 115(2), 224-241.

Weliamto, W. A., Seah, H. S., and Wibowo, A., 2006. Robust Moment Invariant with Higher Discriminant Factor Based on Fisher Discriminant Analysis for Symbol Recognition. *Graphics Recognition. Ten Years Review and Future Perspectives.* 3926, 408-421.

Wen-ting, C., Yong-xue, L., Man-chun, L., Yu, Z., and Zhen, L., 2010. A best-first multivariate decision tree method used for urban land cover classification. *18th International Conference on Geoinformatics.* 1-5.

Wong, W.-H., Siu, W.-C., and Lam, K.-M., 1995. Generation of moment invariants and their uses for character recognition. *Pattern Recognition Letters.* 16(2), 115-123.

Wren, C. R., Azarbayejani, A., Darrell, T., and Pentland, A. P., 1997. Pfinder: real-time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 19(7), 780-785.

Wu, X., Xu, D., Duan, L., and Luo, J., 2011. Action recognition using context and appearance distribution features. *In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 489 –496.

X.Sun, M.Chen, and A.Hauptmann., 2009. Action recognition via local descriptors and holistic features. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR) Workshops . 58-65.

Xiaobin, L., Hui, F., and Jianmin, X., 2008. Automatic Vehicle Classification Based on Video with BP Neural Networks.*4th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM '08)*. 28(6), 205-208.

Xiong, M., Lees, M., Cai, W., Zhou, S., and Low, M. Y. H., D.-H., 1995. A dynamic algorithm for tank objects recognition. *IEEE International Conference on Systems, Man and Cybernetics, 1995. Intelligent Systems for the 21st Century*. 2, 1877-1881.

Xu, D., and Li, H., 2008. Geometric moment invariants. *Pattern Recognition.* 41(1), 240-249.

Yaakob, S. N., and Jain, L., 2012. An insect classification analysis based on shape features using quality threshold ARTMAP and moment invariant. *Applied Intelligence*. 37(1), 12-30.

Yaakob, S. N., 2006. Classification of Binary Insect Images Using Fuzzy and Gaussian Artmap Neural Networks: Kolej Universiti Kejuruteraan Utara Malaysia.

Yang, J. C., Shin, J. W., and Park, D. S., 2007. Fingerprint Matching Using Invariant Moment Features.*Computational Intelligence and Security*.4456, 1029-1038).

Yang, Y., and Webb, G., 2002. A Comparative Study of Discretization Methods for Naive-Bayes Classifiers. Proceedings of the 2002 Pacific Rim Knowledge Acquisition Workshop (PKAW'02), Tokyo, Japan. 159-173.

Yasin, H., and Khan, S. A., 2008. Moment Invariants Based Human Mistrustful and Suspicious Motion Detection, Recognition and Classification. *Proceedings of the Tenth International Conference on Computer Modeling and Simulation.* 734-739.

Yinan, S., Liu, W., and Wang, Y., 2003. United moment invariants for shape discrimination. *Proceedings. IEEE International Conference on Robotics, Intelligent Systems and Signal Processing*. 1, 88-93.

Ying-Han, P., Connie, T., Jin, A. T. B., and Ling, D., 2003. Palmprint authentication with Zernike moment invariants. *Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology (ISSPIT 2003)*. 199-202.

Zhao, D., Shao, L., Zhen, X., and Liu, Y., 2013. Combining appearance and structural features for human action recognition. *Neurocomputing*. 113, 88-96.

Zhao, X., Li, X., Pang, C., and Wang, S., 2013. Human action recognition based on semi-supervised discriminant analysis with global constraint. N*eurocomputing*. 105, 45-50.

Zhengmao, Z., Premaratne, P., Monaragala, R., Bandara, N., and Premaratne, M., 2010. Dynamic hand gesture recognition system using moment invariants. *5th International Conference on Information and Automation for Sustainability (ICIAFs)*. 108-113.

Zhihu, H., and Jinsong, L., 2010. Analysis of Hu's moment invariants on image scaling and rotation. *2nd International Conference on Computer Engineering and Technology (ICCET)*. 7, 476-480.

Zhong, Q., 2008. Method of vehicle classification based on video. *IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM 2008)*. 162-164.

Zhu, Y., De Silva, L. C., and Ko, C. C., 2002. Using moment invariants and HMM in facial expression recognition. *Pattern Recognition Letters*. 23(1–3), 83-91.