CHAPTER – 1

INTRODUCTION

1.1 Activity Recognition

Activity recognition aims to acknowledge the actions and goals of one or a lot of agents from a series of observations on the agent's actions and therefore the environmental conditions. Since the Nineteen Eighties, this analysis field has captured the eyes of many technology communities thanks to its strength in providing personalized support for several totally different applications and its association to several different fields of study like medication, human-computer interaction, or social science.

Associate degree older man wakes up at dawn in his tiny apartment, wherever he stays alone. He lights the stove to form a pot of tea, switches on the kitchen appliance, and takes some bread and jelly from the cabinet. When he takes his morning medication, a computer-generated voice gently reminds him to show off the toaster. Later that day, his girl accesses a secure web site wherever she scans a check-list, which was created by a detector network in her father's lodging. She finds that her father is feeding unremarkably, taking his medication on schedule, and continued to manage his standard of living on his own. That info puts her mind relaxed.

Many different applications are studied by researchers in activity recognition; examples embody helping the sick and disabled. For instance, Pollack et al. show that by mechanically observing human activities, home-based rehabilitation will be provided for folks tormented by traumatic brain injuries (Pollack et al., 2003). One will notice applications starting from security related applications and supply support to location-based services. Thanks to its many-faceted nature, totally different fields could consult with activity recognition as set up recognition, goal recognition, intent recognition, behavior recognition, location estimation and location-based services.

1.2 Types of Activity Recognition

1.2.1 Sensor-Based, Single-User Activity Recognition

Sensor-based activity recognition integrates the rising space of detector networks with novel data processing and machine learning techniques to model a good variation of human activities (Tanzeem Choudhury et al.,2008; Nishkam Ravi et al., 2005). Mobile devices (e.g. good phones) offer ample detector information associated in calculation of power to change physical activity recognition to supply an estimation of the energy consumption throughout way of life. Sensor-based activity recognition researchers believe that by empowering present computers and sensors to observe the behavior of agents (under consent), these computers are higher suited to act on our behalf.

Levels of Sensor-Based Activity Recognition

Sensor-based activity recognition could be a difficult task attributable to the inherent yelling nature of the input. Thus, applied mathematics modeling has been the most thrust during this direction in layers, wherever the popularity at many intermediate levels is conducted and connected. At the bottom level wherever the detector information area unit collected, applied mathematics teaching issues a way to realize the elaborated locations of agents from the received signal information. Associated intermediate level, applied mathematics illation is also involved regarding a way to acknowledge individuals activities from the inferred location sequences and environmental conditions at the lower levels. Moreover, at the very best level a serious concern is to seek out the general goal or sub goals of Associate in provider agent from the activity sequences through a mix of logical and applied mathematics reasoning.

1.2.2 Sensor-Based, Multi-User Activity Recognition

Recognizing activities for multiple users mistreatment on-body sensors, initially appeared within the work by ORL mistreatment active badge systems within the early 90's. Different detector technology like acceleration sensors were used for distinctive cluster activity patterns throughout workplace eventualities. Activities of Multiple Users in intelligent environments are self-addressed in Gu et al. during their work, they investigate the basic downside of recognizing activities for multiple users from detector readings in a very home surroundings, and propose a completely unique pattern mining approach to acknowledge each single-user and multi-user activities in a very unified answer (Tao Gu et al., 2009).

1.2.3 Sensor-Based Cluster Activity Recognition

Recognition of cluster activities is basically totally different from single or multi-user activity recognition therein the goal is to acknowledge the behavior of the cluster as Associate in entity, instead of the activities of the individual members inside it (Dawud Gordon et al. 2013). Cluster behavior is nascent in nature, which means that the properties of the behavior of the cluster are basically totally different then the properties of the behavior of the people inside it, or any add of that behavior. The biggest challenge in modeling the behavior of the individual cluster members is because the roles of the individual inside the cluster is dynamic (Hirano, T et al., 2013) and their relationship to nascent behavior of the cluster in parallel (Bjorn et al., 2007). Challenges that should be self-addressed embody quantification of the behavior and roles of people UN agency be a part of the cluster, integration of express models for role description into illation algorithms, and measurability evaluations for terribly massive teams and crowds. Cluster activity recognition has applications for crowd management and response in emergency things, likewise as for social networking and Quantified Self applications (Dawud Gordon, 2014).

1.2.4 Vision-Based Activity Recognition

It is a awfully necessary and difficult downside to trace and perceive the behavior of agents through videos taken by numerous cameras. The first technique used is pc vision. Vision-based activity recognition has found several applications like human-computer interaction, computer programmed style, mechanism learning, and police work, among others. Scientific conferences wherever vision based mostly activity recognition work typically seem are ICCV and CVPR.

In vision-based activity recognition, an excellent deal of labor has been done. Researchers have tried variety of ways like optical flow, Kalman filtering, Hidden Mathematician Models, etc., underneath totally different modalities like single camera, stereo, and infrared. Additionally, researchers have thought-about multiple aspects on this subject, as well as single pedestrian pursuit, cluster pursuit, and police investigation born objects.

Recently some researchers have used RGBD cameras like Microsoft Kinect to find human activities. Depth cameras add further dimension i.e. depth that traditional 2nd camera fails to produce. Sensory information from these depth cameras are accustomed to generate period of time skeleton model of humans with totally different body positions. These skeleton information provides substantive information that researchers have to be compelled to accustom model human activities that are trained and later accustom acknowledge unknown activities (Piyathilaka L et al. 2013).

1.2.4.1 Levels of Vision-Based Activity Recognition

In vision-based activity recognition, the machine method is usually divided into four steps, specifically human detection, human pursuit, act recognition and a high-level activity analysis.

1.2.4.2 Automatic Gait Recognition

One way to spot specific folks is how they walk. Gait-recognition code may be accustomed to record somebody's gait or gait profile together information for the aim of recognizing that person later, albeit they're carrying a disguise.

1.3 Approaches of Activity Recognition

1.3.1 Activity Recognition Through Logic and Reasoning

Logic-based approaches keep track of all logically consistent explanations of the determined actions. Thus, all potential and consistent plans or goals should be thought-about. Kautz provided a proper theory of arrange recognition (H. Kautz, 1987). He delineates arranged recognition as a logical abstract thought method of restriction. All actions, plans are uniformly brought up as goals, and a recognizer's data is diagrammatic by a collection of first-order statements referred to as event hierarchy encoded in first-order logic, that defines abstraction, decomposition and purposeful relationships between forms of events.

Kautz's general framework for arrange recognition has Associate in exponential time quality in worst case, measured within the size of input hierarchy (H. Kautz, 1987). Lesh and Etzioni went one step more and bestowed ways in scaling up goal recognition to proportion his work computationally (N. Lesh and O. Etzioni,1995). In distinction to Kautz's approach wherever the arrange library is expressly diagrammatic, Lesh and Etzioni's approach allows automatic plan-library construction from domain primitives. What is more, they introduced compact representations and economical algorithms for goal recognition on massive arrange libraries.

Inconsistent plans and goals are repeatedly cropped once new actions arrive. Besides, they conjointly bestowed ways for adapting a goal recognizer to handle individual behavior given a sample of a human recent behavior. Pollack et al. delineate an instantaneous argumentation model that may comprehend the relative strength of many sorts of arguments for belief and intention description (Pollack et al., 2003).

A serious downside of logic-based approaches is their inability or inherent impracticability to represent uncertainty. They provided no mechanism for preferring one consistent approach to a different and was incapable of deciding whether or not one specific arrange is additional possible than another, as long as each of them is consistent enough to elucidate the actions determined. There's conjointly an absence of brainpower related to logic based mostly ways.

Another approach to logic-based activity recognition is to use stream reasoning supported Answer Set Programming, and has been applied to recognizing activities for health-related applications, that uses weak constraints to model a degree of ambiguity/uncertainty.

1.3.2 Activity Recognition Through Probabilistic Reasoning

Probability theory and applied math learning models are additionally applied recently in activity recognition to reason regarding actions, plans and goals.

Plan recognition is done as a method of reasoning below uncertainty, which is convincingly argued by Charniak and Goldman (E. Charniak and R.P. Goldman, 1993). They argued that any model that doesn't incorporate some theory of uncertainty reasoning cannot be adequate. Within the literature, there are many approaches that expressly represent uncertainty in reasoning regarding Associate in Nursing agent's plans and goals.

Using device information as input, Hodges and Pollack designed machine learning-based systems for characteristic people as they perform daily routine activities like creating occasional. Intel analysis (Seattle) workplace and University of Washington have done some vital works on exploitation sensors to find human plans. A number of these works infer user transportation modes from readings of radio-frequency identifiers (RFID) and international positioning systems (GPS) (M.R Hodges and M.E Pollack, 2007).

The use of temporal probabilistic models has been shown to perform well in activity recognition and usually surmount non-temporal models. Generative models like the hidden Markov model (HMM) and also the additional typically developed dynamic Bayesian networks (DBN) are widespread selections in modeling activities from device information. Discriminative models like Conditional Random Fields (CRF) are conjointly ordinarily applied and also provide sensible performance in activity recognition.

Generative and discriminative models each have their pros and cons. Ideal selection depends on their space of application. A dataset along with implementations of variety of widespread models (HMM, CRF) for activity recognition is given.

Conventional temporal probabilistic models like the hidden Markov model (HMM) and conditional random fields (CRF) model directly model the correlations between the activities and also the determined device information. In recent years, increasing proof has supported the employment of graded models that take under consideration the wealthy hierarchical structure information that exists in human activity data. The core plan here is that the model doesn't directly correlate the activities with the device information, however instead breaks the activity into sub-activities (sometimes brought up as actions) and models the underlying correlations consequently. An example can be the activity of getting ready pasta, which might be de-escalated into the sub activities or actions of cutting vegetables, cookery the vegetables during a pan and serving it on a plate. Examples of such a hierarchical model are Layered Hidden Markov Models (LHMMs) (Nuria Oliver et al.2004) and the hierarchical hidden Markov model (HHMM) (TLM van Kasteren et al., 2011), which have been shown to significantly outperform its non-hierarchical counterpart in activity recognition.

1.3.3 Wi-Fi-Based Activity Recognition

When activity recognition is performed inside the cities and in the market Wi-Fi signals and 802.11 access points, suffer a lot of noise and uncertainty. These uncertainties are shapely employing a dynamic Bayesian network model by Yin et al(Jie Yin et al, 2004). A multiple goal model that may reason regarding user's interleaving goals is bestowed by Chai and Yang (Xioyong Chai and Qiang Yang, 2005), where a settled state transition model is applied. a stronger model that models the coincidental and interleaving activities during a probabilistic approach is projected by Hu and Yang (Derek Hu and Qiang Yang, 2008),. A user action discovery model is bestowed by Yin et al., wherever the Wi-Fi signals are segmental to provide potential actions.

An elementary downside in Wi-Fi-based activity recognition is to estimate the user locations. Two vital problems are the way to cut back the human labeling effort and the way to address the dynamic signal profiles once the surroundings changes. Yin et al. restrained the second issue by transferring the labeled data between time periods (Jie Yin et al., 2005). Chai and Yang projected a hidden Markov model-based technique to increase labeled data by investment the untagged user traces (Xioyong Chai and Qiang Yang, 2005). J. Pan et al. (Jeffrey Pan et al. 2007)propose to perform location estimation through on-line co-localization, and S. Pan et al. (Sinno Pan et al.,2007) projected to use multi-view learning for migrating the labeled information to a replacement period.

1.3.4 Data Mining Based Mostly Approach to Activity Recognition

Different from ancient machine learning approaches, Associate in approach supported data processing has been recently projected. Within the work of Gu et al., the downside of activity recognition is developed as a pattern-based classification problem (Tao Gu et al. 2009). They projected an information mining approach supporting discriminative patterns that describe important changes between any two activity categories of knowledge to acknowledge successive, interleaved and coincidental activities during a unified answer. Gilbert et al. use 2nd corners in each area and time (Gilbert et al., 2011). These are classified spatially and temporally employing a graded method, with Associate in increasing search space. At every stage of the hierarchy, the foremost distinctive and descriptive options are learned with efficiency through data processing.

1.4 Image Processing

Each image is scanned into a binary image at a resolution of three hundred dots per inch, when that median filtering is applied to get rid of speckle noise. The image dimensions don't seem to be normalized. Later, the DRT of every image is calculated. Every column of the DRT represents a projection or shadow of the image at an explicit angle. When these projections are processed and normalized, they represent a collection of feature vectors (observation sequence) for the image in question. The DRT of a picture is calculated as follows.

Assume that every image consists of Ψ pixels in total, which the intensity of the ith element is denoted by I_i, $i = 1, \ldots, \Psi$. The DRT is calculated exploitation β non overlapping beams per angle and Θ angles in total. The additive intensity of the pixels that lie among the j^{th} beam is denoted by R_j , $j = 1, \ldots, \beta\Theta$. This can be referred to as the j^{th} beam add. In its distinct type, the atomic number 86 remodels will be expressed.

The accuracy of the DRT is set by Θ (the variety of angles), β (the variety of beams per angle), and also the accuracy of the interpolation technique. Note that the continual variety of the atomic number 86 remodels is inverted through analytical suggests that the DRT so contains virtually a similar data because the original image might be with efficiency calculated with Associate in algorithmic rule by Brace well (R.N. Brace well,1995). Our system calculates the DRT at Θ angles. These angles are equally distributed between 0˚ and 180°.

The dimension of every projection is later altered from β to d. This can be done by initial decimating all the zero-valued parts from every projection. These decimated vectors are then contracted or distended to a length of d through interpolation. Though most the data within the original image is contained within the projections at angles that vary from 0° to 180°, the projections at angles that vary from 180° to 360° also are enclosed within the observation sequence. These further projections are additional to the observation sequence so as to make sure that the sequence fits the topology of our HMM.

Since these projections are merely reflections of the projections already calculated, no further calculations are necessary. Associate in Nursing observation sequence so consists of T = 2Θ feature vectors, that is, XT1 =. Every vector is later normalized by the variance of the intensity of the whole set of T feature vectors. Every image pattern is so diagrammatic by Associate in observation sequence that consists of T observations, wherever every observation could be a feature vector of dimension d.

The DRT, as a feature extraction technique, has many blessings. Though the DRT isn't a shift invariant illustration of an image, shift and scale exchangeability is ensured by the next image process. Every image could be a static image and contains no dynamic data. Since the feature vectors are obtained by scheming projections at totally different angles, simulated time evolution is formed from one feature vector to subsequent, whenever the angle is that the dynamic variable. This allows USA to construct Associate in Nursing HMM for every image.

The DRT is calculated at Associate in which phase vary from 0° to 360° and every observation sequence is then shapely by an HMM of that the states are organized during a ring. This ensures that every set of feature vectors is rotation invariant. Our system is additionally sturdy with regard to moderate levels of noise.

1.5 Images Modeling Using HMM

HMM-based system developed uses endless initial order HMM to represent every image. The HMM-based and DTW-based systems use similar verification protocols A pattern recognition system, that relies on HMMs, usually uses Associate in Nursing HMM to represent every pattern category. Every of those HMMs is employed to model Associate in Nursing observation sequence, similarly because the relationship between the individual observations. HMMs are so created in such how that time-evolution is assumed from one observation within the sequence to subsequent. Since speech signals and dynamic (on-line) pictures conjointly contain temporal data, it's potential to extract endless observation sequence from these signals in a very intuitive method. For this reason HMMs are particularly well-suited for modeling these forms of signals. This can be not the case for static (off-line) pictures. Consequently, feature vectors got to be extracted from off-line pictures in such how that time-evolution is simulated from one observation to subsequent.

In this dissertation use a grid to section a picture into native sq. cells. From every cell, the element density is computed, in order that every element density represents a neighborhood feature. Every image is so diagrammatic by a sequence of feature vectors, wherever every feature vector represents the element densities related to a column of cells. The HMM-based system developed simulates time-evolution from one observation to subsequent by scheming the DRT of every image throughout the feature extraction method. Before we tend to discuss the HMM-based image model, we tend to initial gift the notation within the following section.

The HMM-based system developed in simulates time-evolution from one observation in Associate in observation sequence to subsequent by scheming the DRT of a raw image. The feature vectors are so obtained by scheming projections of a image at totally different angles, when that they're subjected to some more process. The angle is so the dynamic variable. This allows USA to construct Associate in Nursing HMM for every image.

1.6 Human Action and Activity Recognition

1.6.1 Human Activity Recognition Problem Description

Several approaches for human action recognition have been proposed. A survey on HAR can be found at (T. B. Moeslund et al. 2006). A variety of approaches use features which describe the motion and/or shape of the entire human body figure to perform human action recognition. Efros et al. (A. Efros et al., 2003) recognize the actions of small scale figures using features derived from blurred optical flow estimates. Blank et al. (M. Blank et al., 2005) represent an action by considering the shape carved by its silhouette in time.

Local shape descriptors based on the Poisson equations are computed, then aggregated into a global descriptor by computing moments. Another group of methods uses features derived from small-scale patches, usually computed at a set of interest points. Schuldt et al. (C. Schuldt et al..2004) have computed local space time features at location selected in a scale-space representation. These features are used in an SVM classification scheme.

Traditional approaches for motion analysis mainly involve the computation of optical flow (J.L. Barron et al., 1994) or feature tracking (S.M.Smith and J.M. Brady, 1995; A. Blake and M. Isard, 1998). Although very effective for many tasks, both of these techniques have limitations. Optical flow approaches mostly capture first order motion and may fail when the direction of motion has sudden changes. Feature trackers often assume a constant appearance of image patches over time and may hence fail when the appearance changes, for example, in situations when two objects in the image merge or split. Model-based solutions for this problem have been presented by (M.J. Black and A. D. Jepson, 1998).

Image structures in videos are not restricted to constant velocity and/or constant appearance over time. On the contrary, many interesting events in videos are characterized by strong variations in the data along both the spatial and the temporal dimensions. In the spatial domain, points with a significant local variation of image intensities have been extensively investigated in the past (W. Forstner and E. Gulch, 1987; C. Harris and M. Stephens, 1988; T. Lindeberg, 1998; C. Schmid et al., 2000). Such image points are frequently referred to as Interest points and are attractive due to their high information content and relative stability with respect to perspective transformations of the data.

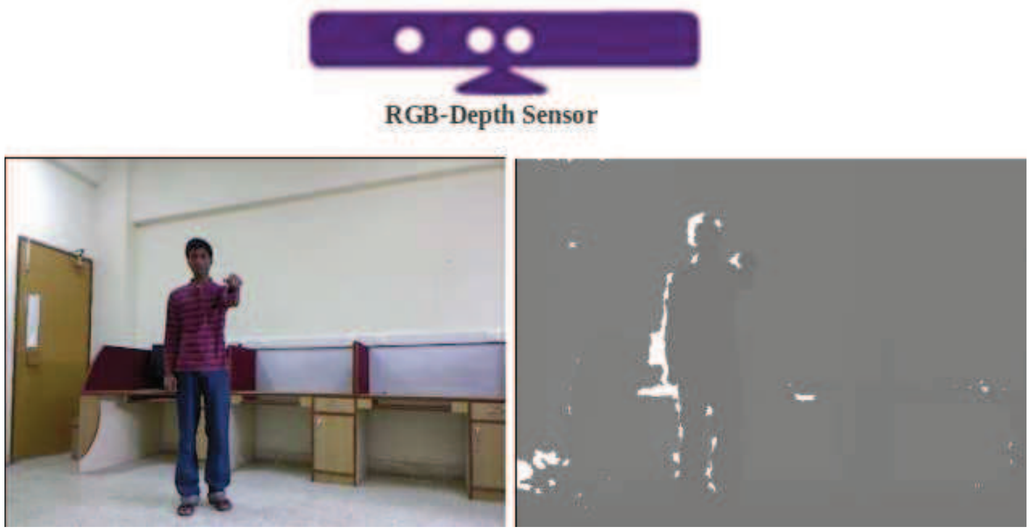

Figure 1.1: RGB and Depth Frame (I. Laptev and T. Lindeberg, 2004)

In this work Human Action Recognition (HAR) for both DEPTH sequence as well as RGB video sequence are analyzed. In depth based HAR-the depth information (Figure 1.1) of video sequence is used whereas in RGB based HAR - the extended notion of interest points (proposed by IVAN LAPTEV-2004) (I. Laptev and T. Lindeberg, 2004) into the spatio-temporal domain is used for a compact representation of video data as well as for interpretation of spatio-temporal events. Latent Dirichlet Allocation (LDA) is then used for modeling the human action.

In principle, activity recognition can be exploited to great societal benefits, especially in real-life, humancentric applications such as eldercare and healthcare. Successful research, however, has so far focused on recognizing simple human activities. Recognizing complex activities remains a challenging and active area of research. Specifically, the nature of human activities poses the following challenges:

- Recognizing concurrent activities. People can do several activities at the same time one such as watching television while talking to friends. These behaviors should be recognized using a different approach from that for sequential activity.

- Recognizing interleaved activities. Certain real-life activities can be interleaved.1 For instance, if a friend calls while you're cooking, you'd talk to your friend for a while, while you continue to cook.
- Ambiguity of interpretation. Similar situations can be interpreted differently. For example, an "open refrigerator" can belong to several activities, such as "cooking" or "cleaning."
- Multiple residents. More than one resident can be present in many environments.

A smart space for example, a smart house needs to recognize the activities being performed in parallel. Human activity understanding encompasses activity recognition and activity pattern discovery. The first focuses on accurate detection of human activities based on a predefined activity model. Therefore, an activity recognition researcher builds a high-level conceptual model first, and then implements the model by building a suitable pervasive system.

On the other hand, activity pattern discovery is more about finding unknown patterns directly from low-level sensor data without any predefined models or assumptions. Thus, an activity pattern discovery researcher builds a pervasive system first and then analyzes the sensor data to discover activity patterns. Although the two techniques are different, both aim to improve human activity technology. In addition, they're complementary to each other—the discovered activity pattern can help define activities that can be later recognized and tracked.

The goal of activity recognition is to recognize common human activities in real-life settings. Accurate activity recognition is challenging because human activity is complex and highly diverse. Researchers have used several probability-based algorithms to build activity models. The hidden Markov model (HMM) and the conditional random field (CRF) are among the most popular modeling techniques.

1.6.2 Expandable Data-Driven Graphical Modeling of Human Actions Based on Salient Postures

Their work presented a graphical model for learning and recognizing human actions. Specifically, they propose to encode actions in a weighted directed graph, referred to as *action graph*, where nodes of the graph represent salient postures that are used to characterize the actions and are shared by all actions (Wanqing Li et al., 2008). The weight between two nodes measures the transitional probability between the two postures represented by the two nodes. An action is encoded as *one or multiple paths* in the action graph. The salient postures are modeled using Gaussian Mixture Models (GMM).

Both the salient postures and action graph are automatically learned from training samples through unsupervised clustering and expectation and maximization (EM) algorithm. The proposed action graph not only performs effective and robust recognition of actions, it can also be expanded efficiently with new actions. An algorithm was also proposed for adding a new action to a trained action graph without compromising the existing action graph. Extensive experiments on widely used and challenging datasets have verified the performance of the proposed methods, its tolerance to noise and viewpoints, its robustness across different subjects and datasets, and as well as the effectiveness of the algorithm for learning new actions.

The last 20 years have seen ever-increasing research activity in the field of human activity recognition. With activity recognition having considerably matured, so has the number of challenges in designing, implementing, and evaluating activity recognition systems. This tutorial aims to provide a comprehensive hands-on introduction for newcomers to the field of human activity recognition. It specifically focuses on activity recognition using on-body inertial sensors. They examined the key research challenges that human activity recognition shares with general pattern recognition and identified those challenges that are specific to human activity recognition. They then describe the concept of an Activity Recognition Chain

(ARC) as a general-purpose framework for designing and evaluating activity recognition systems.

They provided details of each component of the framework, provided references to related research, and introduce the best practice methods developed by the activity recognition research community. They concluded with the educational example problem of recognizing different hand gestures from inertial sensors attached to the upper and lower arm. They illustrated how each component of this framework can be implemented for this specific activity recognition problem and demonstrate how different implementations compare and how they impact overall recognition performance.

Automatic recognition of physical activities-commonly referred to as Human Activity Recognition (HAR)-has emerged as a key research area in Human-Computer Interaction (HCI) and mobile and ubiquitous computing. One goal of activity recognition is to provide information on a user's behavior that allows computing systems to proactively assist users with their tasks (Gregory D .Abowd et al. 1998). Traditionally, research in computer vision has been at the forefront of this work.

A large number of researchers investigated machine recognition of gestures and activities from still images and video in constrained environments or stationary settings (Sushmita Mitra and Tinku Acharya, 2007; P. Turaga et al., 2008 ; J.K. Aggarwal and M.S. Ryoo, 2011). Efforts to recognize activities in unconstrained daily life settings caused a shift toward using inertial sensors worn on the body, such as accelerometers or gyroscopes. Advances in sensor technology now allow for form factors and battery lifetimes suitable for long-term recordings, computing, and continuous interaction on the move.

On-body sensing extends the potential application areas of activity recognition beyond instrumented rooms and promises to provide smart assistance and interfaces virtually anywhere and at any time by observing activities from the user's perspective. At the end of the 1990s, researchers performed the first feasibility studies on activity recognition using body-worn sensors, where the choice of activities seemed arbitrary and not always relevant to real-world applications.

Still, the continuing success of activity recognition motivated steps toward more challenging and application-oriented scenarios.

Several real-world domains were identified that would clearly benefit from activity recognition, such as the industrial sector (Maurtua et al. 2007; G.Ogris Stiefmeier et al. 2008), office scenarios, the sports and entertainment sector (K. Kunze et al. 2006; David Minnen et al. 2006a; Cassim Ladha et al. 2013), and health care. Specifically, the Activities of Daily Living (ADLs) (S.Katz et al. 1970) attracted a great deal of interest (for examples see Ling Bao and Intille, 2004 ; Ravi et al., 2005 ; B.Logan et al., 2007; E.Tapia et al., 2004). Monitoring daily activity to support medical diagnosis, for rehabilitation, or to assist patients with chronic impairments was shown to provide key enhancements to traditional medical methods (T. Starner et al. 1997; M. Sung et al. 2005; J. Chen et al. 2005; N.Oliver and F. Flores-Mangas 2007; Marc Bachlin et al. 2009; Bernd Tessendorf et al. 2011a; Thomos Plotz et al., 2012).

Early assistance to encourage humans to adopt a healthy lifestyle was regarded as another important goal. This led to a vast exploration of related human activities, for example, brushing teeth (Jonathan Lester et al. 2006) or hand washing, food (Oliver Amft et al. 2007; G.Pirkl et al. 2008) and medication intake (D.Wan 1999; R. de Oliveira et al. 2010), or transportation routines (J.Krumm and E. Horvitz 2006). Recently, activity recognition made its debut as a key component in several consumer products. For example, game consoles such as the Nintendo Wii and the Microsoft Kinect rely on the recognition of gestures or even full-body movements to fundamentally change the game experience.

While originally developed for the entertainment sector, these systems have found additional applications, such as for personal fitness training and rehabilitation, and also stimulated new activity recognition research (J. Sung et al. 2011) Finally, some sports products such as the Philips Direct life or the Nike running shoes integrate motion sensors and offer both amateur and professional athletes feedback on their performance. All of these examples underline the significance of human activity recognition in both academia and industry. Despite considerable advances in inferring activities from on-body inertial sensors and in prototyping and deploying activity recognition systems (Bjorn Hartmann et al. 2007; Ashbrook and Starner 2010), developing HAR systems that meet application and user requirements remains a challenging task. This is the case even if HAR techniques that were successfully used for one recognition problem are to be adopted for a new problem domain.

Although activity recognition shares many methodological challenges with other fields, such as computer vision, natural language processing, or speech recognition, it also faces a number of unique challenges and requires a dedicated set of computational methods that extend on those developed in these fields. For example, computer vision and speech recognition can lend themselves to clear problem definitions, such as "detect object in image" or "detect a spoken word in a sentence," and focus on a well-defined and fixed sensing system (i.e., a defined number and type of cameras or microphones).

In contrast, HAR offers more degrees of freedom in terms of system design and implementation (see Table I for a description of the main characteristics of human activity recognition systems). First, there is no common definition, language, or structure of human activities that would allow us to formulate a clear and common problem statement (which activity has to be recognized, how a specific activity is characterized, etc.). For some applications, such as long-term behavioral monitoring, relevant activities can often not even be clearly defined up front. Second, human activity is highly diverse and its recognition therefore requires careful selection of several heterogeneous sensors that differ in their capabilities and characteristics. Sensor composition can also change as sensors may be added and removed opportunistically based on current application requirements (D. Roggen et al. 2009). Finally, activity recognition typically requires specific evaluation metrics to reflect the quality of the system for the intended application.

Activity Recognition (AR), which identifies the activity (e.g., walking, sitting, reading) that a user performs, has generated a great deal of interest within ubiquitous and mobile computing. In particular, the recent explosion of smart mobile devices with sensing, processing, and network capacity have opened up a huge range of possibilities for activity recognition. Given the large numbers of researchers and companies working in this area, one might expect that there would be many deployed activity recognition applications. Surprisingly, however, relatively little practical work has been done on the uses and applications of activity recognition with mobile devices.

In their work they describe and categorize possible activity recognition applications, with the hope that it will encourage the development of such applications and influence the direction of current activity recognition research (Jeffrey W. Lockhart et al., 2012). To illustrate the general process for sensor-based mobile activity recognition, They used Actitracker (Actitracker. https://actitracker.com) application which is built upon the Wireless Sensor Data Mining (WISDM) Platform (Lockhart et al., 2011). Actitracker runs on smart phones and collects readings from the accelerometer sensor. These readings were then transformed into examples which summarize short (10-second) periods with simple features such as average acceleration and frequency. Predictive models were then built using lightweight classification algorithms, such as neural networks and J48 decision trees, which can predict activities like walking or jogging with accuracies above 98% (Kwapisz et al.2012; Weiss et al.2011). These models are used to predict a user's activities throughout the day and the results were made available to the user via a web interface.

Many other activity recognition projects and platforms are being developed (Lane et al., 2010), but generally, there are few practical, deployed applications proposed for AR. Meta-applications such as "Code in the Air" are even being developed to make application development and AR easy for developers, as well as for end users who want to develop their own activity context-aware applications (Ravindranath et al., 2012). However, these efforts leave the question of what to do with AR up to the developers and users, and at this point such platforms are underutilized. With the availability of these platforms, it is even more important to find useful applications for AR. Activity recognition applications fall across a broad range of disciplines. Generally speaking, there are three major types of applications: those that benefit end users, those that benefit developers or third parties, and those that benefit crowds and groups using the application. These types of applications are not mutually exclusive; an application that benefits groups will invariably benefit end users. In the following section we describe a number of activity recognition applications and organize them by categories.

In principle, activity recognition can be exploited to great societal benefits, especially in real-life, human centric applications such as elder care and healthcare (Eunju Kim et al. 2010). Their article focused on recognizing simple human activities. Recognizing complex activities remains a challenging and active area of research and the nature of human activities poses different challenges.

The protection of critical transportation assets and infrastructure is an important topic these days. Transportation assets such as bridges, overpasses, dams and tunnels are vulnerable to attacks (R. Bodor et al, 2003). In addition, facilities such as chemical storage, office complexes and laboratories can become targets. Many of these facilities exist in areas of high pedestrian traffic, making them accessible to attack, while making the monitoring of the facilities difficult. In their research, they developed components of an automated, "smart video" system to track pedestrians and detect situations where people may be in peril, as well as suspicious motion or activities at or near critical transportation assets.

The software tracks individual pedestrians as they pass through the field of vision of the camera, and uses vision algorithms to classify the motion and activities of each pedestrian. The tracking is accomplished through the development of a position and velocity path characteristic for each pedestrian using a Kalman filter. With this information, the system can bring the incident to the attention of human security personnel. In future applications, their system could alert authorities if a pedestrian displays suspicious behavior such as: entering a "secure area," running or moving erratically, loitering or moving against traffic, or dropping a bag or other item.

1.6.3 Action Recognition Based on a Bag of 3D Points.

Their work presents a method to recognize human actions from sequences of depth maps. Specifically, we employ an action graph to model explicitly the dynamics of the actions and a bag of 3D points to characterize a set of salient postures that correspond to the nodes in the action graph (Wanqing Li, 2010). In addition, they proposed a simple, but effective projection based sampling scheme to sample the bag of 3D points from the depth maps. Experimental results have shown that over 90% recognition accuracy was achieved by sampling only about 1% 3D points from the depth maps. Compared to the 2D silhouette based recognition, the recognition errors were halved. In addition, they demonstrate the potential of the bag of point's posture model to deal with occlusions through simulation.

1.6.4 Activity Recognition Using a Combination of Category Components and Local Models for Video Surveillance.

Their work presented a novel approach for automatic recognition of human activities for video surveillance applications (Weiyao Lin et al., 2008). They proposed to represent an activity by a combination of category components and demonstrate that this approach offers flexibility to add new activities to the system and an ability to deal with the problem of building models for activities lacking training data. For improving the recognition accuracy, a confident-frame-based recognition algorithm is also proposed, where the video frames with high confidence for recognizing an activity are used as a specialized local model to help classify the remainder of the video frames. Experimental results show the effectiveness of the proposed approach.

1.6.5 Group Event Detection with a Varying Number of Group Members for Video Surveillance.

Their work presented a novel approach for automatic recognition of group activities for video surveillance applications (Weiyao Lin et al.,2009). They proposed to use a group representative to handle the recognition with a varying number of group members, and use an asynchronous hidden Markov model (AHMM) to model the relationship between people. Furthermore, they propose a group activity detection algorithm which can handle both symmetric and asymmetric group activities, and demonstrated that their approach enables the detection of hierarchical interactions between people. Experimental results show the effectiveness of our approach.

1.7 Problem Definition

Human detection based action recognition is a constantly expanding research area due to number of applications in surveillance (behavior analysis), security (pedestrian detection), control (human-computer interfaces), content based video retrieval, etc. It is, however, a complex and difficult-to-resolve problem because of the enormous differences that exist between individuals, both in the way they move and their physical appearance, view-point and the environment where the action is carried out.

Consequently, we are able to distinguish between similar actions by only considering the body parts which have major contributions to those actions e.g. legs for walking, running etc; hands for boxing, waving etc. Human activity Recognition is a challenging problem in computer vision. One of its main goals is the understanding of the complex human visual system and the knowledge of how humans represent activity in order to discriminate different identities with high accuracy. A general way of defining the problem is : given still images, recognize people at the scene by using a activity database. This problem, taken into account, has several challenges to be solved. Different illumination of the scene; changes in pose, orientation, expression and face occlusions are some examples of the issues to deal with. In our case, the problem is simplified by working under controlled conditions, which can be summarized as follows: one person‘s image, known background, uniform illumination, frontal pose, neutral expression and no occlusions. Two basic and conceptually independent problems have to be addressed by this kind of systems: activity detection and recognition of the detected activity. My work on the recognition stage, is to take the detected activity as the input to the algorithm. This stage can be separated in two steps: feature extraction, where important information for discrimination is saved, and the matching step, where the recognition result is given with the aid of an activity database. For this system, Gabor filter is used for selecting Gabor features for Human Activity recognition. A small subset of Gabor features capable of discriminating from other Human Activity images that are stored in the database. In this dissertation the system developed uses the hidden Markov model (HMM) to match a test Human Activity image with an appropriate reference image. This system uses the Gabor filter to extract the sequence of informative Gabor features from the given Human Activity image. The extracted features are again subjected to discrete Radon transform (DRT) to extract a sequence of feature vectors from a image. The HMM-based system developed in this dissertation matches the feature set (observation sequence) for a test image with an HMM of the claimed image, through Viterbi alignment. A distance measure is obtained by calculating a negative log likelihood.

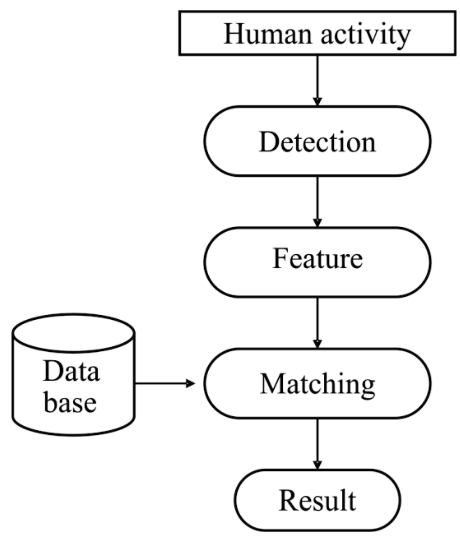

Figure 1.2 Block diagram of proposed system

The system process is divided in to two stages shown in figure 1.2:

1. Image detection.
2. Recognition.

The recognition stage can be separated into two steps:

1. Features extraction: Where important information for discrimination is saved.
2. Matching step: Where recognition result is given with the help of Human Activity database.

In this concept the system developed uses the hidden Markov Model (HMM) to match a test Human Activity image with an appropriate reference image. This system use the Gabor filter to extract the sequence of informative Gabor features from the given Human Activity image. The present study was undertaken with the following objectives:

- To study Gabor Filter and Hidden Markov Model.
- To analysis the recently proposed human activity using Gabor filter with Hidden Markov Model and comparison with exiting forms of Human activity recognition system.
- To study related methods of quality improvement of images.
- To practical implement Gabor Filter transform human activity recognition with Hidden Markov Model through MATLAB simulations.
- To Critical evaluate the original results obtained from individual implementation of existing and novel algorithm for selected application using Gabor Filter with Hidden Markov Model.
- To compare the present result with the result obtained from the previous methods.