DEVELOPMENT OF AN AUTOMATIC ATTITUDE RECOGNITION SYSTEM: A MULTIMODAL ANALYSIS OF VIDEO BLOGS

Noor Alhusna Madzlan

Thesis submitted for the Degree of Doctor in Philosophy
School of Linguistics, Speech & Communication Sciences

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

University of Dublin
Trinity College

Under the Supervision and Direction of
Asst. Prof. Breffni O'Rourke and Prof. Nick Campbell

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

Summary

Communicative content in human communication involves expressivity of socio-affective states. Research in Linguistics, Social Signal Processing and Affective Computing in particular, highlights the importance of affect, emotion and attitudes as sources of information for communicative content. Attitudes, considered as socio-affective states of speakers, are conveyed through a multitude of signals during communication. Understanding the expression of attitudes of speakers is essential for establishing successful communication. Taking the empirical approach to studying attitude expressions, the main objective of this research is to contribute to the development of an automatic attitude classification system through a fusion of multimodal signals expressed by speakers in video blogs. The present study describes a new communicative genre of self-expression through social media: video blogging, which provides opportunities for interlocutors to disseminate information through a myriad of multimodal characteristics. This study describes main features of this novel communication medium and focuses attention to its possible exploitation as a rich source of information for human communication. The dissertation describes manual annotation of attitude expressions from the vlog corpus, multimodal feature analysis and processes for development of an automatic attitude annotation system. An ontology of attitude annotation scheme for speech in video blogs is elaborated and five attitude labels are derived. Prosodic and visual feature extraction procedures are explained in detail. Discussion on processes of developing an automatic attitude classification model includes analysis of automatic prediction of attitude labels using prosodic and visual features through machine-learning methods. This study also elaborates detailed analysis of individual feature contributions and their predictive power to the classification task.

Related Publications

[1] Noor Alhusna Madzlan, Jingguang Han, Francesca Bonin and Nick Campbell. Towards automatic recognition of attitudes: Prosodic analysis of video blogs. *Speech Prosody*. Page 91-94, 2014

[2] Noor Alhusna Madzlan, Jingguang Han, Francesca Bonin and Nick Campbell. Automatic recognition of attitudes in video blogs - Prosodic and visual feature analysis. *Fifteenth Annual Conference of the International Speech Communication Association, INTERSPEECH*. Page 1826 - 1830, 2014

[3] Noor Alhusna Madzlan, Yuyun Huang and Nick Campbell. Automatic classification and prediction of attitudes: Audio-visual analysis of video blogs. *Speech and Computer: 17th International Conference, SPECOM*. Springer. Volume 9319, Page 96 - 104, 2015

[4] Noor Alhusna Madzlan, Justine Reverdy, Francesca Bonin, Loredana Sundberg Cerrato and Nick Campbell. Annotation and Multimodal Perception of Attitudes: A Study on Video Blogs. *Third European Symposium of Multimodal Communication (MMSYM)*. Page 50-54, 2015

Contents

List of Tables

List of Figures

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi
05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

List of Abbreviations

Abbreviation	Meaning	Page
SVM	Support Vector Machine	9
RF	Random Forest	9
VLOG	Video Blog	10
AES	Affective-Epistemic States	18
F0	Fundamental Frequency	19
HNR	Harmonicity-to-Noise ratio	19
B1	Bandwidth of the First Formant	25
H1	Amplitude of the First Harmonic	26
A1	Amplitude of the First Formant	26
H1-H2	First harmonic relative to the Second Harmonic	26
H1-A3	First harmonic relative to the Third Formant	26
NAQ	Normalised Amplitude Quotient	26
ASM	Active Shape Model	28
FAC	Facial Action Coding system	29
AAM	Active Appearance Model	29
DCT	Discreet Cosine Transform	30
HMM	Hidden Markov Model	30
ANOVA	Analysis of Variance	31
HCI	Human Computer Interaction	32
wMEI	Weighted Motion Energy Images	33
GP	Gaussian Process	36
GMM	Gaussian Mixture Model	40
LDA	Linear Discriminant Analysis	41
Vlogger	Video Blogger	47
Hz	Hertz	80
dB	Decibel	80
SEC	Second	80
FMEAN	Mean value of the fundamental frequency	80
FMIN	Minimum value of the fundamental frequency	80
FMAX	Maximum value of the fundamental frequency	80
FPCT	Percentage of shape of the pitch contour	80
FVCD	Percentage of the vibration of the vocal folds	80
PMEAN	Mean value of power/ intensity	80
PMIN	Minimum value of the power/ intensity	80
PMAX	Maximum value of the power/ intensity	80
PPCT	Intensity movement	80
DN	Duration	80
PCA	Principal Component Analysis	90
PC	Principal Components	90
PC1	First Principal Component	97
PC2	Second Principal Component	97
PC6	Sixth Principal Component	97
ASR	Automatic Speech Recognition	100
RBF	Radial Basis Function	114

Chapter 1

Introduction

This dissertation highlights affective expressions as a source of information content in human communication. The study explores the dynamics of attitudinal expressions conveyed through several types of non-verbal signals. Attitudinal expressions between communicators must be understood and displayed appropriately to ensure successful message transfer during the communication process. This research further extends understanding of attitudinal aspects of expression, not only through the exploration of human perception of attitudes, but also through the development of automatic attitude recognition which will be of great utility for automatic understanding of user perception and information retrieval.

This chapter briefly discusses human communication, its different signalling modalities for communicating content of various types, as well as the role of attitudinal states in the communicative setting. Further discussion about the meaning of attitudes and their relation, or non-relation, to affect and emotions is also explored. Several definitions and concepts are introduced to give a clearer overall view of the methodology involved in the study. The final part of this chapter explains the research motivation and objectives.

1.1 Human Communication

Research in domains including social psychology, social signal processing and affective computing has increasingly focused on understanding the dynamics of human communication. Humans communicate with each other in a number of different modalities and the interpretation of their utterances may vary from a person to the other. Generally, humans communicate with each other to share information. Human communication involves a series of processes and can take place between two or more interlocutors. This sharing of information is transferred through signals.

Allwood [1] elaborates on the purpose of human communication, which is dissemination of information. Allwood mentions several types of information involved in the communication process:

1. physiological states (fatigue and hunger)
2. character, identity, personality (being timid, aggressive)
3. affective-epistemic attitudes (showing joy, friendliness, surprise)
4. factual content giving information about beliefs, assumptions about facts
5. communication management (feedback, turn-taking)

Among the information streams used in communication, one interesting type of information is the expression of affective-epistemic attitudes. This type of information is useful for understanding affective states and emotions of the people involved in the communicative setting. Emotions, affect and attitudes seem to be similar concepts but there are clear distinctions between them. Damasio [2] distinguishes between emotion and feelings, stating that feelings refer to the inner, cognitive experience of an emotion while emotion is the observable response of these feelings. The frequently cited five basic emotions are sadness, anger, fear, surprise, disgust and happiness [3]. There are certainly other emotional ways of representing emotional states than the simply categorical, including valence based representations [4] but such concepts are not sufficient to describe the complexities of people's affective states [5]. A person could express a mixture of sometimes contradictory feelings simultaneously. For instance, when a person finds that her friend is leaving the country to pursue her dream

job, she expresses first the feeling of happiness and shifts her expression to sadness. Hence, human emotion, although easily shown, are often complex and confusing to understand or interpret.

Affect, on the other hand, is considered a general term for the inward feelings of the human experience. This concept is a broader representation of feelings where emotion contributes to a large part of the overall definition. While Zanna and Rempel [6] refer to affect as "any thoughts that are infused with strong, weak or no emotion at all", Shouse [7] claims that affect has no relation with feelings and emotion. The term affect is sometimes used interchangeably with emotion. It is believed that there are differences between the two concepts. Emotion refers to the display of feelings while affect is a non-conscious experience. Affect is a pre-personal experience of the speaker that unconsciously affects the consequent feelings and actions of the speaker. Hence, affect is viewed as a broad, abstract concept of the humanly cognitive experience, while emotion refers to the inner feelings of the person. Figure 1.1 summarizes the relationship between affect and emotion.

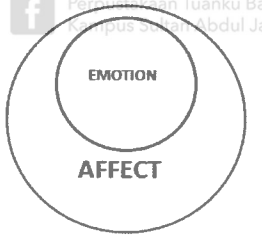

Figure 1.1: Relationship between Affect and Emotion

Affect is a general and broader representation of a person's feelings while emotion is the displayed response of affect, and thus is incorporated into the affective state of a person.

Recent empirical studies, particularly in the field of Affective Computing, use affect to refer to affective-epistemic states of humans when interpreted by machines. This use of affect is a broad term as a point of reference to studies on emotion, attitude and behavioural states of humans displayed in the communicative setting. This study supports the concept of attitude as representation of speakers' affective-epistemic characteristics. Attitude is viewed differently from concepts of affect and emotion. The following section details attitude as a concept for study.

1.2 Attitude as a Communicative Function

The previous section explains the differences between affect and emotion. The concept of attitude is related to affect and emotion, but attitude has distinctive traits. Zanna and Rempel [6] describe attitude as states that may be expressed as strong emotions or may be identified solely from the way an individual behaves with an object. Research on the psychology of attitude finds limited agreement on precise definitions and characteristics of attitude. Psychologists relate attitude with beliefs, opinions, habits and values [8]. Oskamp [8] explains attitude as having three main components, as shown in Table 1.1:

Components	Description
Cognitive	Ideas and beliefs of the agent towards the object
Affective (emotional)	Feelings and emotions the agent has towards the object
Behavioural	The agent's actions towards the object

Table 1.1: Atittude Components

Fishbein and Ajzen [9] suggest that attitude is mainly associated with the affective component. This is believed to be true as these components are independent and separate entities but are still interrelated. The affective component of attitude is an object of interest to several fields of research including psychology, economics, and marketing. However, some researchers use the term attitude loosely and often interchangeably with affect or emotion. In this study, these concepts, although similar, are treated independently. Attitude is of particular interest for this study as it represents observable traits of speaker's cognitive experience. Attitude refers to actions, outer representation of feelings, while emotion refers to inner feelings that are difficult to evaluate. Attitude is believed to be a pragmatic concept for the understanding of affect and emotions [10]. This interpretation of emotional states involves the crucial aspect of intended, voluntary and controlled actions. Auberge [10] explains communicative functions in Figure 1.2.

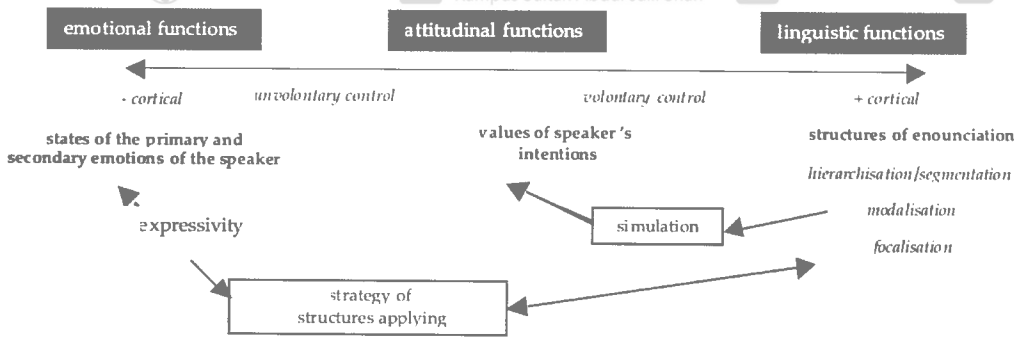

Figure 1.2: Attitude as a Communicative Function

Figure 1.2 shows attitudinal functions as a component for communication. The scale indicates emotional functions to be at the far left of the control scale while linguistic functions are on the far right. Attitudinal function is in the middle of the control scale. This means that attitudes, unlike emotions, are not vague and involuntarily expressed. Rather values of speaker intention are expressed in a conscious, controlled and voluntary manner.

The definition of attitude may or may not have a strong emotional component but always has an evaluative component. In fact, attitude and emotion can be seen as a continuum, a degree of emotional involvement in attitudes. Emotion, conversely, need not have a strong attitudinal - i.e. evaluative - component. This is because emotions are not necessarily directed towards any particular content, that is, you can be happy or sad (for example) without being able to identify why. Additionally, emotions can be internal and less susceptible to conscious control. Whereas attitudes are held more consciously, often deliberately expressed (in language), and subject to control in the sense of reflective evaluation.

Emotion and attitude are conceptually distinct, but interact with one another in reality. Attitudes always have an element of evaluation. Oskamp [8] has identified three components of attitude; cognitive, affective/emotional and behavioural. In this thesis I am concerned with the overt expression of attitude, which I take to be evidence of the first two, internal, components, i.e., the cognitive and affective/emotional states of the speaker. Under one understanding of the term attitude, attitudes can be entirely internal, comprising the cognitive and affective states, but never overtly expressed. For the purposes of this research, the focus is on attitudes as overtly displayed. Therefore, when I use the term attitude, it will frequently refer to the overt expression of attitude.

1.3 Multimodal Communication

Attitude is expressed in the content of communication through several different channels or modalities. Multimodal communication is defined as "co-activation, sharing and co-construction of information simultaneously and subsequently through several modes of perception and production" [11]. This concept adds to the concepts of communication with the sharing of information as its main function. However, multimodal communication involves several modes of information sharing using simultaneous sensory channels. These include sight, hearing, touch, smell and taste. Multimodal communication involves two main processes; firstly multimodal integration (perceiving information based on several modalities) and multimodal distribution (producing information using multimodalities). This is similar to the traditional communicative process, where both communicative agents perceive information and produce feedback. The difference is that there is emphasis on simultaneous sensory modalities.

Multimodal communication has become an important research area as humans perceive and produce communicative expressions through several modalities. A combination of speech, visuals and gesture facilitates communicators to meet their communicative goal of information sharing. As mentioned in Section 1.1, the contents of information include affective-epistemic attitudes. The simultaneous use of multimodal signals facilitates expression of attitudes. Due to the complexities of affective attitudes, the use of more than one sensory modality is beneficial to better understand and interpret attitudes. For example, a person can perceive a friend's expression of friendliness through rising tone of voice, display of a smile and the wave of a hand. Simultaneous use of speech, facial expressions and gestures facilitates the communicators' message transmission.

1.3.1 Multimodalities as Signals in Communication

Numerous studies suggest the relation between display of affective states and non-verbal gestures [12] [13] [14] [15]. Vinciarelli and Valente [16] refer to non-verbal communication as the transmission of a message through non-verbal behavioural cues, such as facial expressions, vocalisations, gesture and posture. Non-verbal communication relays speakers' inner feelings, whether intentionally or unintentionally. For instance, a mother outwardly

expresses approval when her child voluntarily makes the bed by establishing eye contact with the child, praising her with a rising tone and giving a bright smile. In another scenario, the mother tries not to show disapproval when the child throws a tantrum by speaking with a level or falling tone and controlling her facial expressions from showing contempt. This example shows how people communicate feelings and intentions through the dynamics of several and simultaneous non-verbal signals.

With the emergence of the body of knowledge of Social Signal Processing, the role of this area is to firstly provide physical quantification and synthesis of non-verbal signals through which the affective behaviour of humans are expressed, and, secondly, to implement non-verbal signals in conversational agents [12]. This area is especially interesting as multimodal traits of humans are quantified through development of recognisers and synthesisers. This study aims to evaluate and measure communicative contents, in particular attitudinal representation of speakers. The outcome can be helpful in facilitating better understanding of relationships in human-human and human-computer interaction.

1.4 Multimodal Affective Systems

One method of quantifying attitudinal states of speakers through multimodal signals is by applying machine-learning methods. Research in the fields of affective computing, speech technology and human-computer interaction employs machine-learning techniques. There are numerous studies conducted in emotion recognition based on analysis of sentiment [17] [18], prosody [19], facial features [15], gestures [20] and posture [21]. Multimodality (using combination of signals) is also studied in great detail in emotion recognition. However, when a distinct separation of terminology between emotions and attitudes is made, there is little research conducted on automatic attitude recognition. Although there are some notable studies on prosodic attitudes [22] [23][24], research in automatic classification of attitudes using multimodal signals is still scarce.

In this century, there is growing interest in social media as a rich source of human expression. People from different areas of study, including sociolinguistics, psychology, affective computing and human-computer interaction conduct studies using data from social media. Social psychologists, for example, investigate people's communicative purposes in

online blogs [25] [26] [27]. Researchers are also interestered in the social activity of micro-blogging on Facebook [28] [29]. Other empirical studies involve research in online search engines [30] [31], Twitter [32] [33] and YouTube [34] [35] [36]. Through conceptual and empirical research, it is evident that social media offers a rich source of social discovery and scientific application. Similarly, open access to instances of human behaviour, in particular, attitude expressions, from social media users shows a dynamic social representation worthy of exploration and understanding.

1.4.1 Use of Machine Learning Techniques

Machine learning is used in the area of Artificial Intelligence to enable computers to learn without being explicitly programmed. The learning takes place through observations of new data, hence the outcome of this implementation is a system that enables automatic classification of the learned data. Essentially, there are several methods of machine learning – supervised, semi-supervised and unsupervised learning. Supervised learning involves the task of training a dataset that has a group of features and object labels. The machine is constructed to predict and identify the label of an object given the set of features. Unsupervised learning however functions without the identification of labels. Instead, labels are derived based on the training of the dataset given by the features provided.

Classification in supervised learning is used to predict categorical labels given a set of observations. Figure 1.3 [37] illustrates the work flow of building a classification model using supervised learning technique:

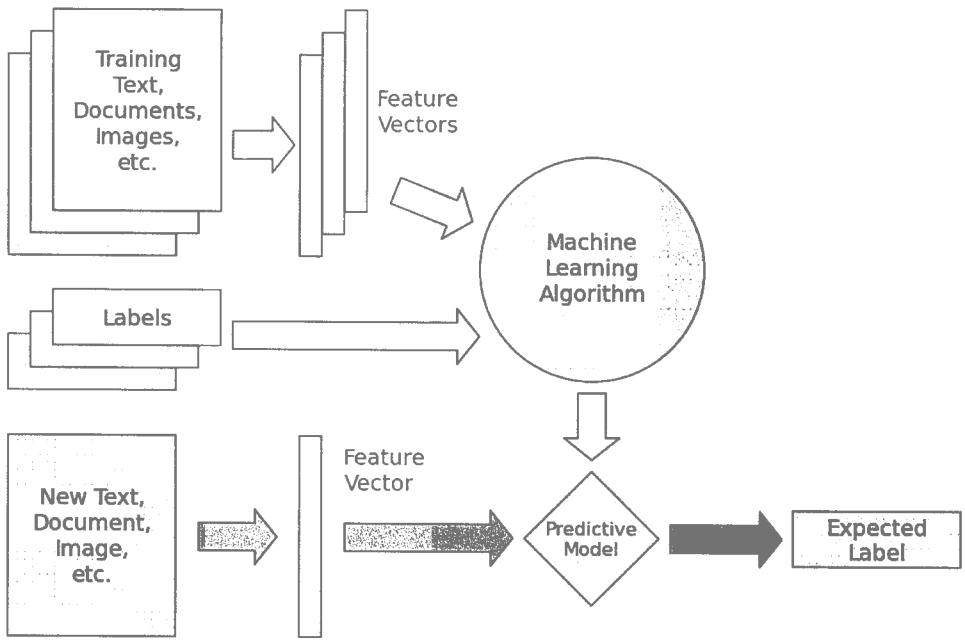

Figure 1.3: Flowchart for Supervised Machine Learning

Figure 1.3 generally describes the processes involved in developing a predictive model using a supervised machine learning method. Following this training process, a predictive model is derived and will produce expected labels.

There are several learning algorithms developed for machine learning classification tasks. One popular method of implementation is the use of statistical learning algorithms [38]. This study adopts this method by using Support Vector Machine (SVM) and Random Forest (RF) classifiers, as will be explained in Chapter 5.

1.5 Conceptual Definitions

This section addresses conceptual definitions used in this thesis. The list describes definitions for the following terminologies:

1. **Annotation**: Annotation is a metadata that is associated to another data by commenting and noting. In this study, annotation involves the act of assigning attitude labels on relevant parts of the video blogs. Annotation of attitude states creates another set of data comprising of only video segments that are annotated with an attitude label.

2. **Attitude Expression**: Attitude refers to pragmatic interpretations of emotion [10]. The expression of attitudes are socio-affective states of speakers expressed in voluntary and controlled settings.

3. **Machine-Learning**: Machine-learning is an area of study in Artificial Intelligence where automatic algorithms are constructed to allow learning and prediction of computers from the given data.

4. **Prosody**: Prosody refers to the suprasegmentals in the voice. The use of prosody in this study covers aspects of fundamental frequency and pitch contours, intensity of the voice, voice quality and duration of speech segments.

5. **Visual**: Visual refers to facial observations from videos used in this study. This reference to visual features includes movement of the eyes and other facial expressions of the speaker's face.

6. **Vlog**: Vlog is an abbreviated term for video blog. This abbreviation is used in this research to refer to YouTube videos of speakers where they share stories of daily life and events.

1.6 Motivation

There has been much work in the fields of Social Signal Processing and Affective Computing in developing recognisers for affective states of speakers through multimodal signals. Human beings can detect differences in affective expressions using data gathered through their eyes and ears. Machines can also do the same using information recorded by cameras and microphones. So by physically capturing and then analysing affective states of speakers, researchers are able to understand human behaviour through concrete displays of affective expressions. Treating attitude as dissimilar to affect and emotion, the development of an attitude recognition system through the means of multimodal signals is an interesting step forward towards understanding human behaviour. Apart from that, development of automatic recognisers has two main purposes:

First Automatic information retrieval and categorisation. Recognition systems are used to index and retrieve information from metadata or other forms of media. Collection of information obtained from these systems enables other systems to implement this source of information for affective modelling.

Second Applications for human-machine interaction. This system could be applied to artificial communicative agents such as social robots for technological advancement. Communicative content can be interpreted by an automatic system and used to inform the behaviour to robots or avatars. This is useful to then develop socially intelligent robots that understand and respond appropriately to humans.

This study aims to contribute to the first purpose of developing automatic recognition systems. Communicative content, in particular, affective information is useful to be embedded into systems for information retrieval and categorisation. This creates a source for people to retrieve information on attitude expressions during speech.

Attitude expressions are especially prevalent in social media where people find comfort and openness in expressing themselves better than in face-to-face communication. One advantage of social media is the ease of public access. This allows researchers to investigate human behaviours, particularly affective expressions from a dynamic source. YouTube, for example, allows public access for people to share videos about their daily life. This rich source of naturally-occurring communication is especially interesting to further understand

human behaviour. Treating these videos as representations of people's attitude expressions makes the collection of this corpus interesting for development of an attitude recognition system that could be applicable to other areas of study.

1.7 Statement of Problem

To the best of my knowledge, little research has so far been conducted in the area of automatic attitude recognition, particularly in exploiting social media corpus such as YouTube video blogs to understand multimodal expressions of humans. This study introduces the concept of attitudes as an evaluative representation of human social experience through spontaneous talk in vlog speech. Following that, internalising attitude states in a computer system for artificial intelligence also contributes to this study's major work by applying multimodal signals into the recognition system.

1.8 Research Objectives

The main objective of this research is to develop an automatic attitude recognition system for information retrieval and categorisation. To achieve this, the following objectives are defined:

Objective 1 Collection of a corpus that represents attitude expressions. Expressions of attitudes are observed and investigated from social media. The use of social media is essential as they provide dynamic and rich multimodal signals that include speech, facial expressions and gestures to indicate speakers' attitude expressions. This study annotates and segments five attitude expressions from vlogs in YouTube.

Objective 2 Investigation of the use of multimodal features as contributors to developing a reliable attitude recognition system. Relevant prosodic and visual features are extracted and analysed. Subsequently, specific prosodic and visual features are selected and examined to understand their contribution towards the classification model.

Objective 3 Development of a reliable automatic attitude recognition system. This research highlights the development of an attitude recognition system that can predict different attitude states of speakers. The development of this system creates a concrete understanding of

attitude expressions through multimodal signals, namely prosodic and visual features. Supervised machine learning is conducted using Support Vector Machines (SVM) and Random Forests (RF).

1.9 Detailed Outline

This thesis is divided into six main chapters which are listed below:

Chapter 1 Discussion of general concepts of human communication and multimodalities in affective expression of attitudes. This chapter also outlines research motivation, statement of problems and three main research objectives for the present study. Conceptual definitions used in this thesis are also listed in this chapter.

Chapter 2 Elaboration of discussion and criticisms of theories of attitudes, multimodalities and use of social media for understanding human communication. This chapter also outlines some of the recent works in the study of automatic recognition through multimodalities using machine learning techniques.

Chapter 3 Discussion of a collection of the vlog corpus. Vlogs are collected from an online video sharing website, YouTube. An ontology of attitude annotation scheme is introduced in this chapter. The processes of collection, annotation and segmentation of attitude states of speakers are described in detail.

Chapter 4 Discussion of the processes involved in developing an automatic attitude classification system. This process involves steps for multimodal feature extraction, from prosodic and visual means. The second part of the chapter discusses processes of feature selection. This section highlights prominent prosodic and visual features that provide greatest contribution to the classification system in recognising different attitudinal states of speakers.

Chapter 5 Supervised machine-learning techniques and results on building a reliable attitude classification system. A total of three experiments are reported:

a) development of an attitude classification system through prosodic modality

b) use of prosodic and visual modalities to develop an automatic attitude classification model

c) improved use of multimodal feature sets for improving the classifier

Chapter 6 Statement of general conclusions, research limitations, suggestions for improvement and future direction of this study.

Chapter 2

State of the Art

This dissertation focuses on developing an attitude recognition system for the purpose of classification and retrieval of human attitudes from video recordings. Prior to more elaborate discussion, this chapter presents the state-of-the-art from related literature addressing the notion of attitude, its relevance to multimodal and novel forms of media, and recent applications in recognition interfaces. The present work adopts some of the conceptual frameworks and methods used from past literature to address the research goals. This chapter begins with a description of related literature concerning attitude definitions and their relation to communicative content.

2.1 Communicative Content

As we have noted, Allwood [1] sees the exchange of information between communicators. We will now explore the five types of communicative information that he identifies (repeated here for convenience):

1. Physiological states
2. Character, Identity and Personality
3. Affective-Epistemic Attitudes
4. Factual Contents
5. Communication / Feedback Management

Physiological states are expressed in the communicative setting, and these include fatigue and hunger. The character of a person is also communicated in speech, examples include the aggressiveness and openness of an individual. Affective-epistemic attitudes refer to the affective or emotive state of the speaker: for instance friendliness, joy and surprise. Examples of factual content include information about beliefs, theories and assumptions, while communication management consists of turn-taking, sequences of feedback and topic change [1]. Another aspect of communicative content that is relevant to the present work in this thesis is Affective-Epistemic Attitudes. This content is essential for the communicative process during information transfer. The next section describes relevant literature relating to attitudes.

2.1.1 Attitude - Definition

Communicative content during communication involves several types of information. One crucial source of information, as mentioned by Allwood [1] is the affective-epistemic attitudinal states of communicators. To better understand the communicative process as well as the modes of interaction, it is essential to first understand the role that attitudes have [1]. In the literature, attitudes are interchangeably related to affect and emotions. However, this thesis identifies attitudes as a distinctive term and this is described through the definition and components of attitudes.

Malhotra [39] states distinctions between affect and attitude, which agrees with Ajzen that attitude refers to summary evaluations of an object or behaviour [9] [40]. Bargh and Chartrand explain attitude as being judgments that are outcomes of spontaneous and unconscious effort [41]. They claim that attitude can be automatically activated without any prior goals of judgment [41]. This perspective, however, is not supported by Auberge [10] who views attitude as driven by intention, voluntary and controlled action. Auberge describes a continuum (see Figure 1.2) of controlled expression of affect where emotion is regarded as an inward and involuntary cognitive experience. Attitude is pragmatic representations of emotions, voluntary and controlled expressions of the cognitive state while linguistic function is the linguistic means of affective expression. Hence, attitude stems from a pragmatic level of conscious and controlled action, and not from a spontaneous and unconscious effort.