# COMPUTERIZED SALES ASSISTANTS: THE APPLICATION OF COMPUTER TECHNOLOGY TO MEASURE CONSUMER INTEREST – A CONCEPTUAL FRAMEWORK

Gurvinder Singh Shergill
Department of Commerce,
Massey University,
Auckland, New Zealand,
G.S.Shergill@Massey.ac.nz

Abdolhossein Sarrafzadeh
Institute of Information and Mathematical Sciences,
Massey University,
Auckland, New Zealand,
h.a.sarrafzadeh@massey.ac.nz

Olaf Diegel
Creative Industries Research Institute,
Auckland University of Technology,
Auckland, New Zealand
olaf.diegel@aut.ac.nz

Aruna Shekar
Institute of Technology and Engineering,
Massey University,
Auckland, New Zealand,
A.Shekar@massey.ac.nz

## ABSTRACT

This paper describes a computerized intelligent sales assistant that gives sales personnel the ability to allocate their time where it will produce the best results, both for the customer, and for the business. Upon entering the shop, a potential customer has his features scanned and analyzed by the computer and the customer is categorized as a browser, future customer, potential customer or buyer. The customers' facial data are also used to retrieve their details, if available, from the shop's database, and the data are used to determine whether a human sales assistant is required. The intelligent assistant's expression recognition feature would also tell the sales personnel whether or not the customer requires or desires assistance in the first place. The paper also proposes a scenario whereby the system could be used to give online sales systems the ability to automatically tailor the services they offer to the customer based on their facial reactions. While browsing the contents of an e-shop, a customer has his facial expressions scanned and analyzed by the computer and, based on the results, the system can suggest further products that may be of interest to the customer. This, for the customer, can mean being directed to products they have been recognized to be more interested in, resulting in savings in time. The framework described in this paper could also be used for applications such as new product screening, marketing and advertising. This paper describes the theoretical and conceptual framework for such an intelligent sales assistant and discusses the technology used in its implementation.

Keywords: feature recognition, facial expression, shopping behavior, consumer analysis and prediction.

## 1. Introduction

Human communication is a combination of both verbal and nonverbal interactions. Through facial expressions, body gestures and other non-verbal cues, a human can communicate with others. This is especially true in the communications of emotions. In fact, studies have shown that a staggering 93% of affective communication takes place either non-verbally or para-linguistically [Mehrabian, 1971] through facial expressions, gestures, or vocal inflections [Picard, 1998]. As discussed in this paper, when dealing with online shoppers, who are communicating with a computer to potentially make a purchase, detection of emotions can certainly play an important role. The same system has the capability to be used in physical stores as well.

Using computers to analyze human faces has been an area of recent interest in computer science and psychology. Smart rooms with computer systems capable of tracking people and recognizing faces, and which can interpret speech, facial expressions and gestures made by the individuals in the smart room have been proposed before [Pentland, 1996]. According to Pentland, the philosophy behind this technology is that "…computers must be able to see and hear what we do before they can prove truly helpful. What is more, they must be able to recognize who we are and, much as another person or even a dog would, make sense of what we are thinking." [p.68]. Robots that show emotions are being built [Bartlett, et al., 2003] while other similar attempts to include emotion in various computer applications [Sarrafzadeh et al., 2002] are being made. Facial recognition software have even be used to apprehend potential terrorists by using the software to scan a crowd of people going through an airport and looking

for facial matches to known terrorists [Kopel and Krause, 2002]. Despite all this, today's computers do not make use of the nonverbal cues common in human communications.

Affective reactions of human subjects in e-commerce environments have been studied [Preira, 2000]. In social psychology researchers have been interested in relating facial expressions with the behavior of a person. So far in marketing there have not been any studies aimed at relating facial expressions with the shopping behavior of customers, and the great potential of such technologies in e-commerce applications have been overlooked. The objective of this paper is exploratory & conceptual and it attempts to relate how the facial expressions of a person might be used to distinguish potential customers from window shoppers when shopping in a physical or online store. Though many people enter a physical or online store, not all end up buying the goods or services from the store. This is especially true about large department stores. The same applies to online shoppers who may browse through shopping sites without actually making purchases. It also describes how, in an e-commerce context, such facial expression recognition could be used to tailor suggestions, or even advertising, to suit the customers requirements.

People visiting a store can be broadly classified into four categories [Stevens, 1989; Moe, 2003]: The Browsers, who just enjoy wandering through the store and killing their own time and the time of sales personnel. Secondly, we have the Future Customer who may be collecting information about a product or service in order to make a future purchasing decision. Thirdly, there are the Potential Customers, who may have the desire to buy a product but are not handled properly by the sales person and hence do not mature a deal with the store. The fourth category is that of the Buyers who are there specifically to purchase a product straight away. It is not always possible for all the sales personnel to distinguish which of these categories the potential customers fit into and thus design selling strategies tailored to those categories in terms of allocation of time to such customers. There are generally two reasons for sales people failing to make the distinction between different categories of customers: Either they do not possess enough training or skills to identify them, or they do not have enough time at their disposal to attend to those customers on whom it is worth spending the time to mature a deal.

We have proved that it is possible to automatically detect facial expressions by developing facial expression analysis software that is capable of detecting six different facial expressions [Sarrafzadeh et al., 2004; Fan et al., 2005]. Based on this proof, we are proposing an intelligent sales assistant that can automatically scan the store and guide sales staff towards potential buyers and suggest suitable sales strategies. Such a tool would be a great asset for large department stores and will enhance sales while decreasing costs. The same is true for online sales activities.

With on-line sales activities, there are currently systems available that will suggest products to customers based on their previous purchases, or even on their browsing behavior. An intelligent facial expression recognition system could provide further depth to such systems by allowing the computer to suggest further products to the customer based on the expressions detected while the customer is browsing through the product catalogue. In many e-commerce situations it is also possible to compare several products to see how they differ in features. In such a scenario, facial expression could also be very useful in detecting which products the customer has an unspoken preference for. The customer is therefore taken to products they have been recognized to be more interested in. The ability of the system to direct the customer to the right products is a desirable feature and results in considerable savings in time as well as reduced frustration. This is a very important issue when one is searching an online shop for the right product.

The facial expression recognition feature of the sales assistant is a general tool for detecting facial expressions. In the context of online sales the facial expression recognition system works well as there is only a single person the system is dealing with and the system has a nearly frontal view of the face. Although meant for use within the sales assistant, this feature of the system has many potential business and other applications such as targeted advertising and marketing both on line and otherwise. Among other applications are assessment of audience mood in television debates and security applications.

## 2. Background Literature

Although there are various studies on the applications of computers in areas related to our proposed intelligent sales systems in the form of e-commerce and m-commerce [Okazaki, 2005], personalized online product selection [Srikumar & Bhasker, 2004], and web-based shopping systems [Kim & Galliers, 2004] we have, so far, found no conceptual or empirical studies in this field. This paper is conceptual in nature and therefore aimed at exploring and initiating the debate in this field.

It has been well researched that facial expressions do reflect cognitive behavior, and that individuals observe other's facial expressions and then use these to regulate their own behavior in social interactions [Salovey & Mayer, 1990]. We do find some studies relating facial expression to some aspects of marketing but none on facial expression and shopping behavior. Howard and Gengler [2001] found that favorable facial expressions do have a positive bias on consumer product attitudes. Sirakaya and Sonmez [2000] used facial expression to study the gender

images in Government tourism brochures. Derbaix [1995] also used facial expression to investigate the effect of television advertisement on attitude towards advertisement and brand attitude. Yuasa et al. [2001] developed a computer network negotiation support tool using facial expression to negotiate an agreement strategy between seller and buyer. They argue that, if players select a happy face, then there is a greater chance they will reach an agreement. Lee [2002] developed computer software programs to implement an e-shopping authentication scheme that used facial recognition instead of user name and password to control access to the web site in a modern consumer e-shopping environment. Consumers would not need to remember their user name and password to get access to the web site. The computer, based upon the facial features of the user, would recognize faces and allow access to the web site to authorized users.

A recent study by Ghijsen et al. [2005] in the context of computer based tutoring systems has shed some light on the types and frequencies of facial expressions exhibited by students during interactions with a computer. These are the same as the results that were found in another recent study of human tutoring sessions [Alexander et al., 2005]; we might tentatively suggest that the expressions displayed by students are not significantly affected by whether the tutor is human or artificial. This would concur with the Media Equation of Reeves and Nass [1996] which argues that interactions between computers and humans are inherently social, and that the same rules for interactions between humans also apply to interactions between humans and computers. However, much more complete studies would be required before firm conclusions could be reached on various aspects of the facial expressiveness of customers.

The authors of this article have developed hardware and software to analyze different facial expressions and then relate these to the behavior of a person in an educational setting [Sarrafzadeh et al., 2003]. The outcome of this research has some useful applications in the field of marketing, especially in sales (online or physical) and the training of sales people. There is also an added advantage that, if sales people can quickly identify the potential customers from window shoppers with the intervention of the system, then they can spend more time on potential customers and convert these potential customers into buyers. It will save the sales people time and the customer will feel well attended to. This will also reduce the expenses on sales forces as the stores will employ less people as they can identify the potential customer and devote sufficient time on them to convert them into buyers instead of spending unreasonable time on window shoppers. Using the same technology, online shoppers can be directed to appropriate products which could result in increased sales.

## 3. Objectives and Research Methodology

This is a multi-staged project and in the first stage (this paper) the objective is to develop a theoretical and conceptual model to initiate a discussion in this area. The next stage will be to use in-house software developed by the research team that is capable of analyzing facial expressions to relate these different facial expressions to purchasers' behavior while they shop. This will enable the system to differentiate between real and window shoppers. The system can then give advice to sales staff based on this information. One of the strategies that will be followed is to give window shoppers freedom to browse without disturbance from sales staff, while real shoppers will have assistance available as required. We believe that this will result in cost savings and perhaps increased sales for store owners.

In the second stage of this research (which is outside of the scope of this paper) the research plan is to use an in-house camera and the computer software developed by the team to classify the different facial expressions of purchasers while they shop. Before we proceed to use the in-house camera and software to classify the shoppers into four categories of shoppers, we will test our categorizations of shoppers. We intend to use a panel of 20 sales personnel working in durable goods stores as respondents to verify our categorization scheme. Firstly, respondents will be asked if they agree or disagree with our four types of shoppers' categorizations. Secondly, the respondents will be shown a number of customer facial expressions while they shop around and asked to classify them into categories. After empirically testing our categorization scheme, a number of face images will be used to train the neural network based software to enable it to classify store visitors into categories of shoppers.

In addition to being able to interpret facial expressions, the software developed has face recognition abilities. We intend to use this capability to identify past shoppers and relate that to previous purchases made by the same person. The system can then be extended to include shopping behavior and preferences of customers in the processing. Store databases will contain sales data and images of customers.

As stated above, in the second stage of research, a number of face images will be used to train the neural network based software in order to enable it to classify store visitors into the four categories of shoppers discussed earlier. We then intend to empirically test the system on real test subjects.

The third phase of this project is to develop software based on the facial expressions of shoppers that can be used as a tool for training sales staff.

Once these three stages of research are finished, the e-commerce stage will begin in which the system will be further developed to use facial expression to detect product preferences in customers which will then make it beneficial in suggesting specific products to customers based on their expressions when browsing through other products. Such a system would be beneficial for both traditional and on-line shopping experiences.

The same system will then also be tested in new product development scenarios in which several product ideas can be screened by customers, and the preferred option can be chosen based on the customers facial expressions. Customers often have unspoken needs or desires that it may be possible to pick up through changes in their facial expressions.

Though this paper deals primarily with a computerized in-store sales assistant, future development will include e-commerce applications such as those described above. In summary, this paper deals with the first stage of the research which is conceptual in nature.

## 4. Significance of facial expressions and Internal State Estimation in Sales and Marketing

In this study, we are using facial expression analysis to estimate the internal state of shoppers and to use that data to identify real shoppers. This is done to save time and to direct sales staff to points where a sales is more likely and to give window shoppers more freedom to browse undisturbed thus increasing the potential for future sales. There is disagreement on the extent to which facial expressions reflect the human internal state [Azar, 2000]. We, however, believe that although facial expressions may not always be the true reflection of emotions, they are one of a number of possible indicators of internal state. Facial expressions may be used as a way of letting others know how one feels about something and can be intentionally deceptive. Hill and Craig [2002] found that they could differentiate between real and faked facial expressions in patients responding to examinations of pain conditions, based on the duration and frequency of the facial expressions.

There is also evidence that the affective state influences processing strategy. Schwarz and Clore [1996] summarize the evidence that during decision-making happy individuals tend to choose top-down strategies, relying primarily on internalized pre-existing knowledge structures, whereas individuals who are sad tend to utilize bottom-up strategies that rely on whatever is being presented at the time of the decision. Since it is relatively easy to detect a happy versus sad affective state through facial expression, the sales staff can then be advised of the strategy that will be the most effective.

### 4.1. Alternative Methods of Detecting Affective State

Although vision based affective state detection is the least intrusive and most applicable way of detecting affect that is possible with the current state of technology there is an alternative. Wearable emotion detection technology [el Kaliouby et al., 2007] can be used for detecting affect. This technology would however require one to attach various devices to the body. Among these devices are data gloves which are very expensive and not feasible for applications similar to the sales assistant application described in this paper. Other devices used for affective state detection purposes include heart rate monitors and pressure sensors. We have developed an intelligent mouse device that is capable of detecting the user's heart rate while they use the system. It is currently unclear how helpful this device will be in a sales situation but this is an area that is being considered for further study.

## 5. Computer-aided Detection of a Shopper's Intent to Purchase

Given that facial expressions and body gestures are a significant factor in identifying real shoppers, how can we use a computer to detect the affective feedback volunteered by the shopper? Current research focuses on facial expressions as perhaps the most important medium of non-verbal, affective communication. That spontaneous (or unconscious) facial expressions really illustrate an affective state is supported by several recent studies that found a high correlation between facial expressions and self-reported emotions [Rosenberg & Ekman, 1994; Ruch, 1995]. Physiological activity has been found to accompany various spontaneous facial expressions [Ekman & Davidson, 1993] which also bear witness to a link between emotions and facial expressions. We can conclude that both spontaneous facial expressions and their accompanying physiological activity illustrate an underlying affective state.

Computers can interpret facial expressions of emotion with reasonable accuracy due to the recent development of automated facial expression analysis systems. Automated facial expression analysis systems identify the motion of muscles in the face by comparing several images of a given subject, or by using neural networks to learn the appearance of particular muscular contractions [Fasel & Luettin, 2003]. This builds on the classic work of Ekman and Friesen, who developed the Facial Action Coding System for describing the movement of muscles in the face [Ekman & Friesen, 1978]. An affective state can be inferred from analyzing the facial actions that are detected [Pantic & Rothkrantz, 1999]. In the following sections we will describe in-house facial recognition/expression analysis software which we intend to discuss in order to theoretically conceptualize how this work will be used in further stages of research.

## 6. An Automatic Facial Recognition/Expression Analysis System

Recent research in computer vision has led to the development of many algorithms for face detection, face recognition, facial feature extraction, and some attempts at automatic facial expression analysis have been made. The goal of this work is the analysis of people's facial expressions from a facial image using a combination and extension of the existing algorithms. Algorithms for face detection, feature extraction and face recognition were integrated and extended to develop a facial expression analysis system that has the potential to be used in intelligent tutoring systems. The new facial expression analysis system uses a fuzzy approach to interpretation of facial expressions, which is another contribution of this paper. Although recognizing people's facial expression is a challenging task, we think it can lead to many exciting developments in human-computer interaction.

Our in-house facial expression recognition system was recently tested with 62 subjects. The subjects interacted with the computer for 20 minutes during which they were videoed and their facial expressions were extracted from the video stream. The facial expression recognition system time-stamped each detected expression, and these were later validated through the videos. The systems performed well in effectively recognizing the facial expressions of the participants. Recognition was performed in real-time in less than 1 second.

This section of the article describes a facial expression analysis system developed in-house. The software uses advanced artificial intelligence techniques including neural networks and fuzzy logic.

### 6.1. Face Detection

The task of face detection is to find the location of a face in an image. It is a necessary first-step in facial expression recognition systems with the purpose of localizing and extracting the face region from an image with often complex and unpredictable backgrounds. It also has several applications in areas such as face recognition systems, crowd surveillance, and intelligent human-computer interfaces. Unfortunately, the human face is a dynamic object and has a high degree of variability in its appearance which makes face detection a difficult problem in computer vision. A wide variety of techniques have been proposed such as color based, motion based, and Artificial Neural Network based face detection.

The color-based approach is the easiest method to perform, by labeling each pixel according to its similarity to skin color, and subsequently labeling each sub-region as a face if it contains a large blob of skin color pixels. Because the head has circular properties we can check circularity to locate the biggest blob as the head. It can deal with different viewpoints of faces, but it is sensitive to skin color and the face shape. It will not work if the background is too similar to the face color or lighting changes, and it can't work from gray-level images. Motion based methods use motion segmentation to locate a face. Neural network methods use a training set of face and non-face images to build a classifier. We used a Neural Network based approach trained using the Carnegie Mellon and ACADEMIA SINICA IIS Face databases. Figure 1 shows training set examples from the databases.
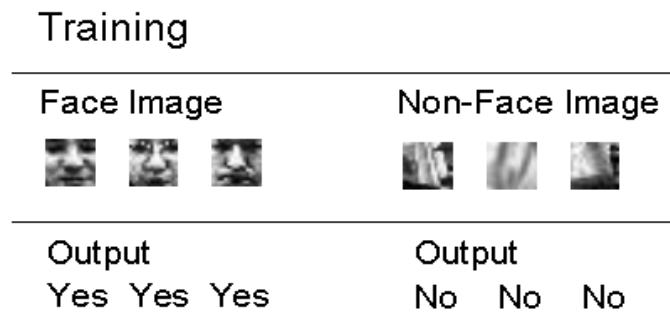


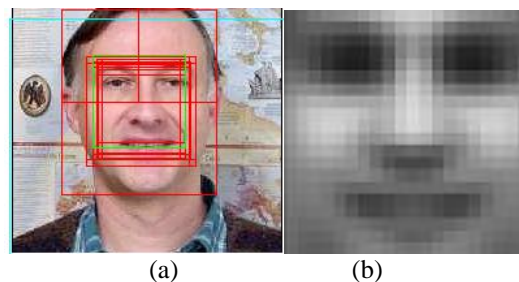Figure 1: Training examples from the face databases.



(a)            (b)

Figure 2: Result of Face Detection and Face Template using a test dataset

Figure 2 (a) shows the result of face detection. The rectangles in the centre of the image show all the possible face locations, the grey rectangle shows the best of these. Figure 2 (b) shows a face template. A face template can be calculated by averaging all the faces in the face database and is used to find the best of the possible face locations.
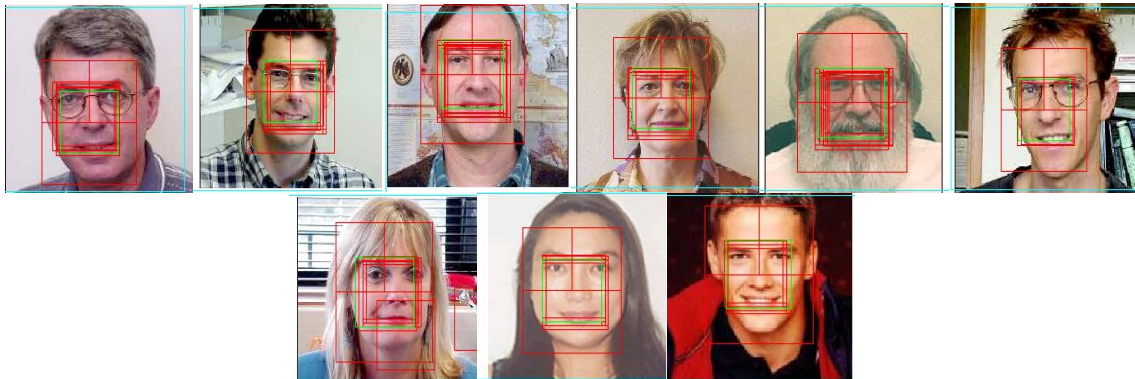


Figure 3: More Results of Face Detection

Figure 3 shows results for a wider variety of faces. In practice we found this method was robust enough for our application.

6.2. Face Recognition

Face recognition is an active area of research with many widespread applications. Humans have a remarkable ability to recognize faces even when confronted by a broad variety of facial geometries, expressions, head poses, and lighting conditions. Research has approached the problem of facial representation for recognition using mathematical techniques. Principal components analysis (PCA) has been a popular technique in facial image recognition. This method addresses single-factor variations in image formation. Thus the conventional 'eigenfaces' facial image recognition technique works best when a person's identity is the only factor that is permitted to vary. PCA is a useful statistical technique that has found applications in fields such as face recognition and image compression, and is a common technique for finding patterns in data of high dimension. A method for computing eigenvectors necessary for PCA was proposed by Turk and Pentland [1991]. This efficient eigenface algorithm has been widely accepted and we applied it successfully to this application.

6.3. Facial Feature Location

We used regions of interest to divide the face into three small regions as shown in Figure 4. The regions are the left eye, the right eye, and the mouth. The face image was divided into these three regions because it made facial extraction local. Inside the eye region we can locate the position of the iris and eyebrow. Inside the mouth region we can locate the mouth. In this way a lot of unnecessary image information is ignored. All of the algorithms for eye detection, mouth detection and Active Contour Modeling algorithms were applied inside the regions of interest.

Dividing the face image into three regions of interest (ROI) is straightforward. As said above these three ROI regions are the left eye's ROI, the right eye's ROI, and the mouth's ROI. The width of the rectangle was the width of the located face multiplied by 1.5, and the height of the blue rectangle was the height of the located face multiplied by 2. The center of the green square rectangle is an approximation to the nose position which also implies the center of the face.
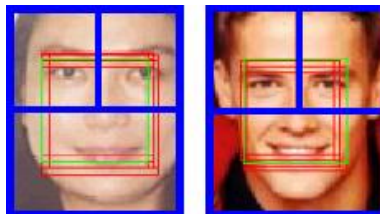


Figure 4: Regions of Interest

In a similar manner to face detection, a Neural Network was used to detect eyes. An eye database was extracted from the face database by hand. Figure 5 shows some sample eyes used for training the Neural Network. The eyes were normalized by scaling to be the same size with the iris in the center of the eye image.
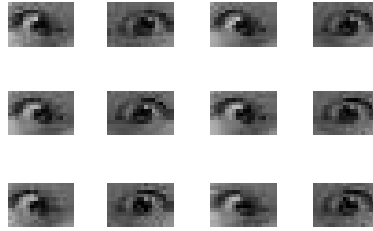
Figure 5: Eye Samples

The process of eye searching is similar to face detection. Figure 6 shows results of eye detection. The grey rectangles show eye locations.

Mouth detection is more complex because mouth shape is very difficult to model. To locate the mouth the image was de-noised first using a wavelet transform to enhance the mouth's contrast. Using edge detection the approximate mouth shape was extracted. One interesting attribute of the mouth is that it must be a closed shape. The mouth must be the biggest blob in the mouth region. Furthermore, if people have a moustache the mouth region and moustache region can be separated as they can be distinguished by comparing the distance between eyes with those two regions. The mouth must be further than the moustache from the eyes. Figure 6 also shows some examples of mouth location.
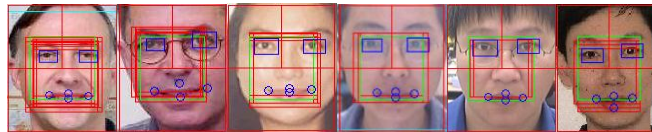


Figure 6: Result of eye and mouth location

6.4. Facial Outline Extraction

Finding an object's contour in an image is a very difficult problem. Kass et al. [1988] represent image contours in a form that allows higher-level processes. This active contour model is defined by an energy function, and a solution is found using the techniques of variation calculus. They proposed an algorithm for the active contour model using dynamic programming. This approach is very stable and allows the inclusion of hard constraints in addition to the soft constraints inherent in the formulation of the function. However, it is slow having complexity of $O(nm^3)$, where n is the number of points in the contour and m is the size of the neighborhood in which a point can move during a single iteration. Williams and Shah [1992] summarize the strengths and weakness of the previous approaches and present a greedy algorithm which has performance comparable to the dynamic programming and variation calculus approaches with a complexity of $O(nm^3)$.
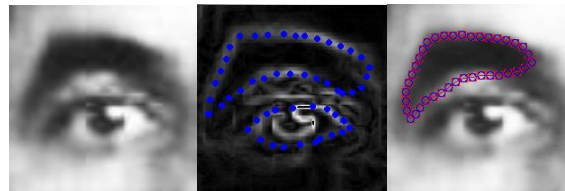


Figure 7: Eyebrow outline extraction using active contours.

Although active contour modeling is a good algorithm to extract an object's outline, it can't generate initial points (by snake itself). The snake point's initialization algorithm is depicted in figure 7. This shows how the outline of the eyebrow is extracted. Figure 8 shows how the same algorithm is applied to the mouth.
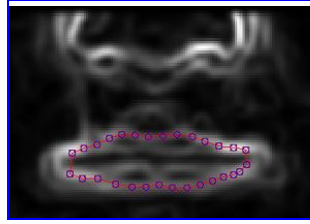
Figure 8: Mouth outline extraction.

6.5 .Facial Analysis Using Fuzzy Logic

After identification of the facial feature outlines, parameters from these were classified using a fuzzy classifier. The classifier was trained using a database of different facial expression samples. A sample of the database images is shown in Figure 9.
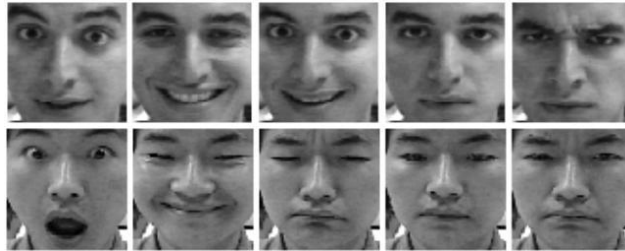


Figure 9: Facial Expression Samples

6.6. Eyebrow Analysis

Facial feature points for the eyebrow may produce many different shapes. However, only reasonable shapes are considered. Five standard eyebrow shapes are defined as shown in Figure 10. Each point is normalized in the range –0.5 to 0.5.

The output indicates shape values in the range -0.5 to +0.5. The system works by mapping 2D shapes to linear space.
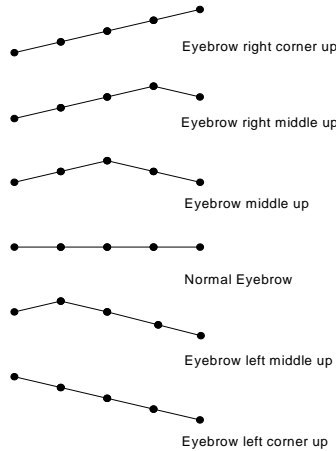


Figure 10:  Basic Shapes of Eyebrow

6.7. Mouth Analysis

Mouth analysis is very difficult because a human mouth may have many different shapes. To simplify this problem we separated the mouth shapes problem into 'Opened mouth shapes' and 'Closed mouth shapes' by testing the mouth feature's points. Figure 11 describes the mouth analysis procedure.
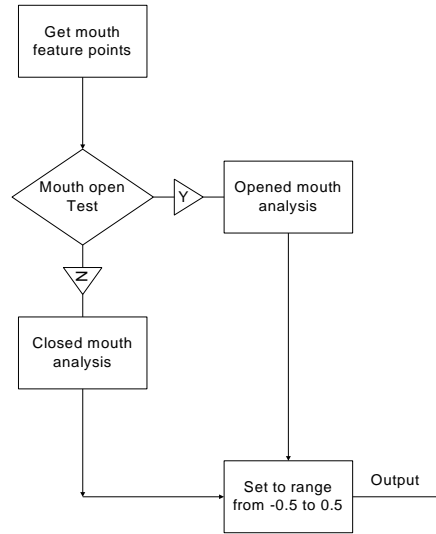
Figure 11:  Mouth Analysis Procedures

Mouth feature points are extracted using a middle line algorithm. The input is a total of 22 points marked using circles in figure 12. The calculation is similar to that of the eyebrow. Fuzzy mouth input points are calculated by comparing the feature points and max length 'a' and max height 'b'.
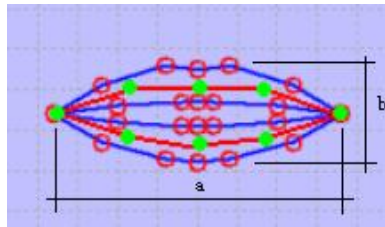


Figure 12:  Mouth Model

Because only some mouth shapes can produce meaningful facial expressions, we defined some of the most important mouth shapes. These shapes can also be mapped to a linear space. We defined six standard mouth shapes, three for an open mouth and three for a closed mouth. Figure 13 shows these shapes. The output is in the range 0.0 to 1.0.
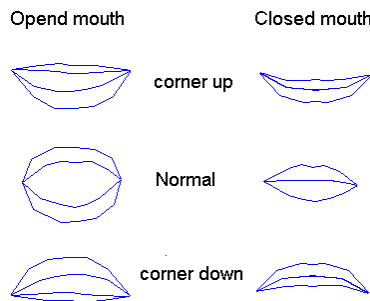


Figure 13: Basic Shapes of the Mouth

The facial feature linear mapping algorithm makes the complex facial expression analysis problem much simpler. Left eyebrow, right eyebrow, eyes and mouth are the most important facial features. These features are to be used as an input for the final decision unit. In the future it is possible to add more facial features such as forehead and moustache.

For the purpose of illustration, four different images were presented to the facial expression analysis system and the system correctly detected the facial expression from the image. Figures 14-17 show examples of the final expression analysis software in use. The algorithm was tested using a small test set. The bars in the boxes to the right of the images show the results. The seven expressions that are detected by the system and the system's certainty in the detection are shown in the box. These certainty values are between 0 and 1 and are calculated by the fuzzy classifier. The longest bar is the best value for the classifier. Images captured via a webcam are analyzed



Figure 14: Normal Expression Detected by the Facial Expression Analysis Software
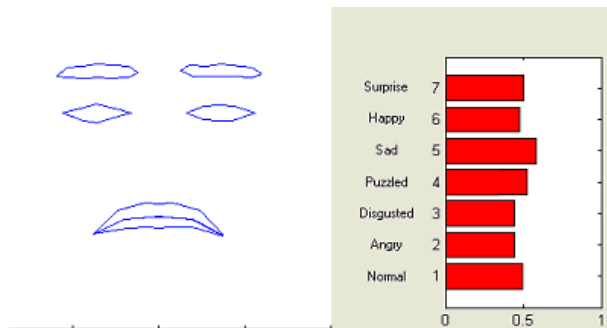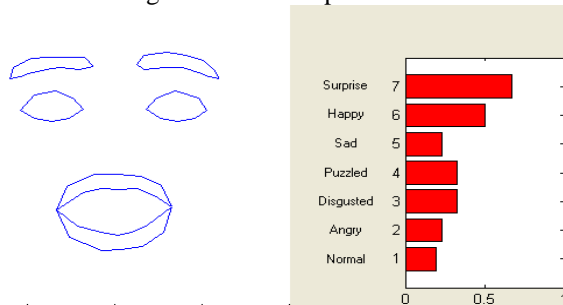


Figure 15: Sad Expression
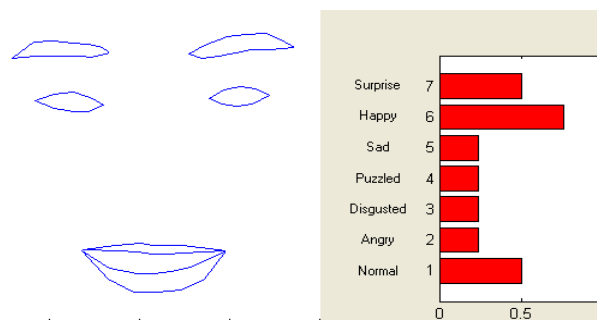


Figure 16: Surprised Expression



Figure 17: Happy Expression

6.8. Testing and results using an improved facial expression recognition component

We recently improved our system using a new approach to facial expression recognition and obtained very promising results. The new approach is based on support vector machines (SVM). A support vector machine is a supervised learning algorithm based on statistical regression introduced by Vapnik [1995]. The SVM algorithm operates by mapping the training set into a high-dimensional feature space, and separating positive and negative samples [Cristianini and Shawe-Taylor, 2000].

The facial expression database used is represented by connection features. Each raw image is 200 pixels in width and 200 pixels in height. By connection extraction, we reduced the image size to 50 pixels by 50 pixels, which is 2500 connection features. This makes the training process significantly faster than the pixel-wise analysis of the image.

An SVM was trained using summarized facial images using our connection features algorithm. The training was obviously considerably shorter using this approach. Memory requirements were also considerably lower. To train SVM, a 5-fold cross-validation was applied. Different Kernel methods are used in SVM to classify non-linear functions. We tested different kernel models, namely a linear model, a polynomial model and an RBF Kernel with the SVM, and have presented the results in Fan et al. [2005]. The training image database contains 1000 image for each facial expression making up 6000 images in total. The results obtained using an RBF kernel were the most promising.

Table 1. The result of applying different kernel models and correct detection for each facial expression

|  | Linear Kernel Model | Polynomial Kernel Model | RBF Kernel Model |
|---|---|---|---|
| Normal | 89% | 85% | 92% |
| Disgust | 78% | 82% | 93% |
| Fear | 83% | 86% | 90% |
| Smile | 91% | 92% | 93% |
| Laugh | 85% | 92% | 96% |
| Surprised | 87% | 93% | 94% |

Video sequences are recorded and analyzed at a rate of 12 frames per second in real time. Our SVM analyzer is triggered each 0.0833 seconds. These results indicate that the proposed intelligent sales assistant is capable of operating in real-time and with a high rate of accuracy. This means that the sales assistant can easily be used in e-commerce applications and online sales and marketing.

## 7. An Intelligent In-Store Sales Assistant

As an intelligent sales assistant, the software package must give sales people as good an indication as possible of what kind of shopper a person entering their store is. This can be determined through the combination of a number of factors.

The first of these factors is in knowing who the customer is, which can be determined through the face recognition portion of the software comparing the face against a database of stored faces. Although faces are not linked to names, it may be argued that ethical issues must be resolved before the use of such technology is permitted. Indeed ethical issues are of prime importance in this context and require extensive studies. Such studies may be as simple as studying how displaying a sign informing customers of the use of such technology or more involved and complex involving law and legislation which is beyond the scope of this study.

Once the face is recognized, the person's past shopping history can be recalled and analyzed. Certain heuristic decisions can then be made depending on the available data. If, for example, the person has a good shopping history at the store, then it may be worth paying them more attention (based on this factor alone). If, however, they are browsing through the washing-machine section, and the database shows them as having purchased a washing machine a few weeks previously, then an intelligent decision might be for the sales person not to spend so much time on them at this moment in time. Although the system can still operate with limited information and without the need to store data, the more data one has available, the better these decisions can become.

The next factor is the recognition of facial expression. When a person is walking past a row of toasters they would normally fit into one of the four potential customer categories outlined in the introduction: The person is not in the least interested in a purchase, they are collecting data for a future purchase, they need convincing by a sales person to make the purchase, or they know exactly what they want to purchase. The last three customer categories are those that may require help from a sales assistant, but each customer category may require help to a greater or lesser degree.

For the potential shopper, who just requires help to make up their mind, the facial expression recognition would be invaluable. In browsing through the many available toasters, for example, the software would determine which model brings the greatest delight to the customer, and the sales assistant would therefore know which model might be the easiest to promote.
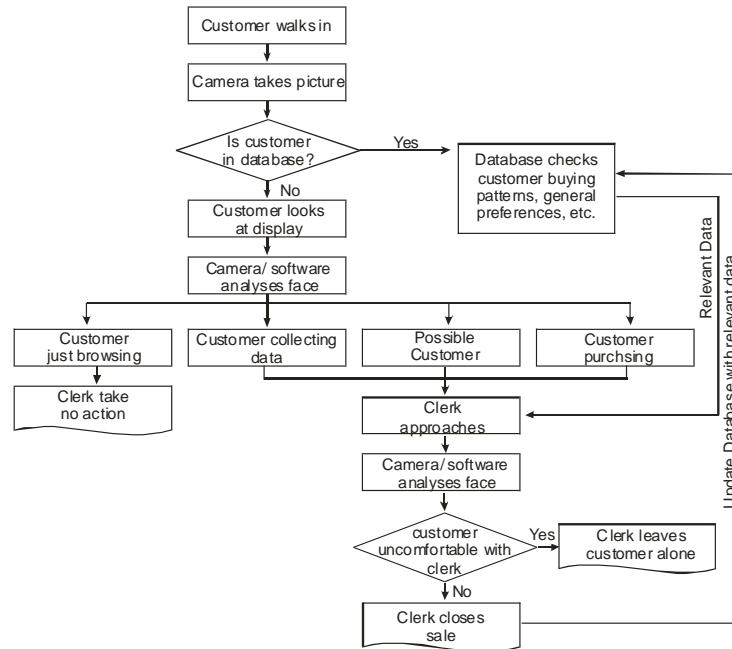


Figure 18: Intelligent In-Store Sales assistant flow diagram

The third factor is that there may also be much value to be gained from knowing a potential customer's reaction to being approached by a sales assistant. Many shoppers, in fact, like to do their browsing in private and get irritated when approached by a sales assistant, as they then feel a sudden pressure to purchase. If the software were able to tell the sales assistant that their presence is making the potential shopper uncomfortable, they could then back off for a while, thus allowing the customer to make their own decision. As the software package would also detect the 'decision made' expression, the sales assistant could then approach the customer and conclude the transaction.

Any one of the above factors alone may lead to erroneous decisions being made by sales staff but the combination of all three factors paints a relatively clear picture of shopper potential and intentions thus allowing sales staff to more effectively spend their time where it is of greatest value.

A block diagram of the Intelligent Sales Assistant Software package would look something like Figure 18.

## 8. An Intelligent On-Line Sales Assistant

Though the initial thrust of this paper is in the development of an in-store intelligent sales-assistant, future development will include the development of an on-line intelligent sales assistant for e-commerce purposes. An example of the use of such an intelligent sales assistant would be in suggesting suitable buying choices to customers as they are browsing through a site. Current 'directed' sales strategies on on-line commerce sites such as Amazon, for example, include directed suggestions based both on a visitors buying history, and on their browsing patterns. If for example, they spend much time browsing the sports section of the site, the computer makes sports related suggestions to the customer.

The addition of a facial expression recognition feature would make such a sales assistant even more powerful. As the buyer now scans through the selection of sports books, for example, the software scans their faces, through the customers webcam, and looks for expressions of varying interest and, based on these, is able to better direct suggestions of other books the customer may be interested in.

A block diagram of an On-Line Intelligent Sales Assistant Software package might look something like the following:
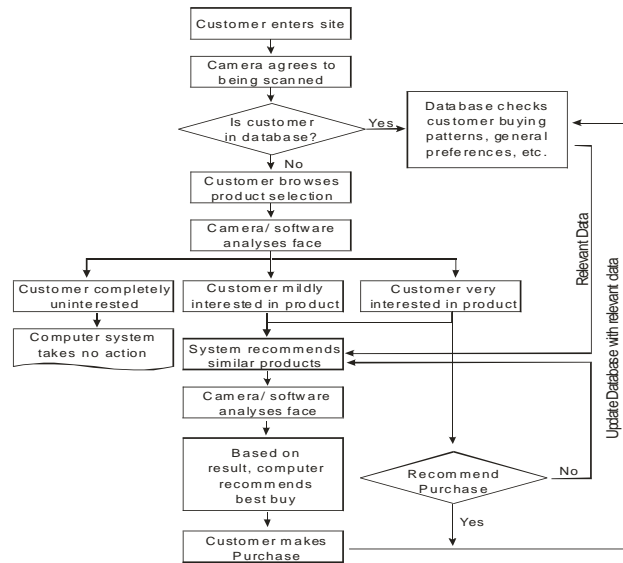
Figure 19: Intelligent On-Line Sales assistant flow diagram

## 9. Ethical Considerations

Though there are obvious ethical considerations in such a scenario that fall outside the scope of this paper, at the very least, there would need to be a stringent process of customer agreement to having their faces scanned through their webcams or shop cameras, and assurances that the data would only be used for the immediate purpose of making purchase recommendations.

It is important to note that such ethical considerations are already present in current customer habit tracking, even though this tracking is not of a visual nature. It currently simply tracks their habits in terms of frequency of visits to different parts of the site, or by frequency of purchase The addition of facial expression recognition simply gives a more precise tool for measuring not just the habits, but also possible explanations (through facial expressions) about those habits.

With the many ethical considerations in the use of such tools, the simplest initial workable solution may be for the customer to have the choice, upon entering a web portal, to go to the conventional section of the site, or to the web camera enabled section of the site. The web camera section of the site would also need a range of explanations as to the use of the system and, most importantly, the use of the data.

Some of the ethical considerations faced by web-enabled online shopping systems are not dissimilar to the use of camera systems as used in many British towns, and in airports around the world. The main ethical dilemma is in the use of the information that is gathered, and this dilemma is bound to become more prevalent with the proliferation of webcams in a variety of areas. These ethical considerations are the same for an online shopping system as they are for an in-store shopping system, other than the methods used for questioning the users about their agreement to the use of such a system.

With the proliferation of video capable mobile phones and Internet communication systems, such as Skype, which now include video, social norms with respect to the use of webcams is likely to evolve considerably in the future. Further research will then be required to see whether users, for example, offer different facial expressions when they consciously know they are being filmed, to those when the filming is unconscious.

## 10. Discussion and Concluding Remarks

The primary question that springs to mind when discussing facial expression recognition is the accuracy and repeatability of such a system. As people grow up and are educated in infinitely varying environments, every individual reacts differently to various stimuli, and an intelligent computerized sales assistant must be able to take these distinctions into account. Questions that we intend to answer through this study include those about the differences in facial expressions between different races, cultures and ages.

The system discussed here has the potential to increase sales both in a store or an online context. Using such a tool can create great changes in the way we do business, advertise and market products. The potential is particularly large in electronic sales and marketing.

Though it is of little doubt that an intelligent computerized sales assistant could have immense value to sales and marketing organizations there are, of course, certain ethical factors that must be considered. One cannot definitely say that the ethical considerations engendered by such a system are either right or wrong, but one must, at the very least take them into consideration, so that any negative impacts can be mitigated. Brown and Muchira [2004] have investigated the issues of privacy and online purchaser behavior and have discussed the implications. Although there are similarities, the issue of privacy in the context of the proposed intelligent sales assistants is one that deserves an investigation in its own right.

Recent literature has highlighted the need for ethical approaches, as new and sophisticated retail customer observation techniques emerge [Kirkup and Carrigan, 2000]. The value of the camera observation techniques lie in their ability to provide accurate and non-disruptive accounts of consumer behavior, but its covert nature raises ethical concerns. There has been an increase in surveillance technology used by retailers and researchers. Generally consumers accept many aspects of information collection and surveillance as a part of modern life but, as they become more aware of the extent to which their behavior is of interest to retailers, concerns are being raised [Kirkup and Carrigan 2000]. Often these can be solved by providing information on the purpose of research and to enhance service to customers and improve the efficiency and effectiveness of retail operations. The research application helps retailers understand consumer behavior, leading ultimately to improvements in the shopping experience. The current code for surveillance research provides general ethical guidance for the conduct of research, and principles from which to develop specific ethically-based research practice. Notices in the retail store could inform customers of the purpose of recording and how they can opt to have their image deleted from tapes. Most codes require that information should not be held longer than is necessary for the purpose for which it is to be used, and provision should be made to destroy data after set times. Access to this data should be controlled. Video surveillance in retail stores is currently legal and unlikely to present a serious threat to consumers. It is important for researchers interested in using surveillance as a research tool to keep abreast of the potential ethical concerns and develop appropriate strategies, such as those outlined here, to reassure and protect consumers.

The system described in this paper assumes that it deals with customers on a one by one basis. This means that only one person will be facing the camera at any one time. In future extensions of the system, the intelligent sales assistant will be able to analyze the facial expressions of groups, and make a decision based on a number of measures including averaging the affect detected. Using parallel processing techniques the system will still be able to perform in real-time. The team is already working on extending the facial expression recognition system for integration into a television debate interface system where the averaged mood of an audience is detected and relayed back to the debatees.

The use of vision based affect detection technology in the sales assistant has several advantages. The first is that this technology compared to alternative methods is far less intrusive and does not require the user to attach devices to their body. The second is that it is cheaper as it does not require additional specialized hardware. The system can be used with a normal processor and very cheap cameras like ordinary webcams.

Though this paper discusses facial expression recognition, it should be remembered that this forms only a small subset of the many areas of human behavior analysis, which also includes topics such as body language, voice tone analysis, and more. Ultimately, the perfect computerized shopping assistant may need to include some, or all, of these areas of research to be successful.

## 11. Future Directions

As stated above, the next stage of this research is to empirically use the concept developed in this paper, which means that facial expression recognition technology will be used to categorize customers into four different customer types. Apart from that, the facial expression recognition technology has enormous applications in areas ranging from elderly healthcare and smart house technologies, through to market research and recruitment. In addition to facial expressions, body movements and gestures can be used as another channel for detecting affective state. We have already developed novel algorithms and software tools [Dadgostar et al., 2005; Dadgostar and Sarrafzadeh, 2006] that will allow us to include gesture recognition into the intelligent sales assistant. And will continue to develop in this direction.

As a future study direction, the facial expression recognition technology could be integrated into the existing Massey University smart house and used to monitor the well-being of the elderly smart house occupants. It could also be used in concept selection with consumers in the new product development process. These applications offer exciting new research opportunities for the future.

**Acknowledgement**

## REFERENCE

Alexander, S.T.V., S. Hill, A. Sarrafzadeh, How do Human Tutors Adapt to Affective State? Proceedings of the Workshop on Adapting the Interaction Style to Affective Factors at User Modeling 2005, Edinburgh, Scotland.

Azar, B. , "What is in a Face?" *Monitor on Psychology*, Vol. 31, No. 1: 91-92, 2000.

Bartlett, M. S., Littlewort, G., Fasel, I. and Movellan, J. R., "Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction", IEEE Conference on Computer Vision and Pattern Recognition, June 16 - 22, 2003, Madison, Wisconsin, 2003.

Brown, M. and Muchira, R., "Investigating the relationship between Internet privacy concerns and online purchase behaviour", *Journal of Electronic Commerce Research*, Vol. 5, No. 1: 62-70, 2004.

Cristianini, N. and Shawe-Taylor, J. , *An Introduction to Support Vector Machines and other Kernel-based Learning Methods*. Cambridge University Press, Cambridge, UK, 2000.

Dadgostar, F. and Sarrafzadeh, A. An Adaptive Real-Time Skin Detector Based on Hue Thresholding: A Comparison on Two Motion Tracking Methods, *Pattern Recognition Letters,* Vol. 27: 1342-1352, 2006.

Dadgostar, F., Sarrafzadeh, A., and Gholamhosseini, H., A Component-based Architecture for Vision-based Gesture Recognition, Accepted in the *Image and Vision Computing New Zealand Conference,* University of Otago, Dunedin, 28 - 29 Nov, 2005.

Derbaix, C. M. "The Impact of Affective Reactions on Attitudes toward the Advertisement and the Brand: A Step toward Ecological Validity", *Journal of Marketing Research*, Vol. 32 (November): 470-479, 1995.

Ekman, P. and Davidson, R.J., "Voluntary smiling changes regional brain activity", *Psychological Science* V. 4, Issue 5: 342-345, 1993.

Ekman, P. and Friesen, W.V., *Facial Action Coding System*. Consulting Psychologists Press, 1978.

Fan, C., Sarrafzadeh, A., Dadgostar, F. and Gholamhosseini, H., "Facial Expression Analysis by Support Vector Regression", *Proceedings of the Image and Vision Computing New Zealand Conference,* University of Otago, Dunedin, 28 - 29 Nov, 2005.

Fasel, B. and Luettin, J., "Automatic Facial Expression Analysis: A Survey", *Pattern Recognition*, Vol. 36, No. 1: 259-275, 2003.

Ghijsen, M., Heylen, D., Nijholt, A., and op den Akker, R., Facial affect displays during tutoring sessions. Affective Interactions: The Computer in the Affective Loop (Workshop at IUI 2005) (San Diego, Ca., 2005).

Hill, M.L. and Craig, K.D., "Detecting deception in pain expression: the structure of genuine and deceptive facial displays", *Pain*, Vol. 98, No. 2: 135-144, 2002.

Howard, D.J. and Gengler, C., "Emotional Contagion effects on Product Attitudes", *Journal of Consumer Research*, Vol. 28 (Sept): 189-201, 2001.

el Kaliouby, R., Picard, R. W., Teeters, A., Goodwin, M., Social-Emotional Technologies For ASD, "International Meeting for Autism Research" Seattle, Washington, May 2007.

Kass, M., Witkin, A., and Terzopoulos, D., "Snakes: Active Contour Models", *International Journal of Computer Vision*, Vol. 1, No. 4: 321-331, 1998.

Kim, C. & Galliers, R. D., "Deriving a diffusion framework and research agenda for web-based shopping systems", *Journal of Electronic Commerce Research*, Vol. 5, No. 3: 199-215, 2004.

Kirkup, M., Carrigan, M., "Video surveillance research in retailing: ethical issues", *International Journal of Retail & Distribution Management*, Vol. 28, No. 11: 470 - 480, 2000.

Kopel, D. and Kruase, M., "Face the facts, Facial Recognition Technology's troubled past and troubling future", http://reason.com/0210/fe.dk.face.shtml, *ReasononLine*, (October), 2002.

Lee, R.S.T., "iJADE Authenticator- An Intelligent Multiagent based Facial Authentication System", *International Journal of Pattern Recognition and Artificial Intelligence,* Vol.16, No. 4: 481-500, 2002.

Mehrabian, A., "*Silent Messages",* Wadsworth, Belmont, California, 1971.

Moe, W.W., "Buying, Searching, or Browsing: Differentiating between Online Shoppers using In-Store Navigational Clickstream", *Journal of Consumer Psychology*, Vol. 13, No. 1&2: 29-39, 2003.

Okazaki, S., "New perspectives on m-commerce research", *Journal of Electronic Commerce Research*, Vol. 6, No. 3: 160-164, 2005.

Pantic, M. and Rothkrantz, L.J.M., "An Expert System for Multiple Emotional Classification of Facial Expressions", *IEEE International Conference on Tools with Artificial Intelligence*, Chicago, USA, 8-10 Nov, 1999.

Pentland, A., "Smart rooms", *Scientific American,* Vol. 274, No. 4: 68-76, 1996.

Picard, R. W., "Towards Agents that Recognize Emotion", *IMAGINA*, Actes proceedings, pp. 153-155, Monaco, 1998.

Preira, R. E., "Optimizing human-computer interaction for the electronic commerce environment", *Journal of Electronic Commerce Research*, Vol., No. 1: 23-44, 2000.

Reeves, B., and Nass, C. I. The Media Equation: How People Treat Computers, Television and New Media Like Real People and Places. Cambridge University Press, 1996.

Rosenberg, E. L. and Ekman, P., "Coherence between expressive and experiential systems in emotion", *Cognition and Emotion*, Vol. 8: 201-229, 1994.

Ruch, W., "Will the real relationship between facial expression and affective experience please stand up: the case of exhilaration?" *Cognition and Emotion,* Vol. 9: 33-58, 1995.

Salovey, P, and Mayer, J., "Emotional Intelligence", *Imagination, Cognitive and Personality,* Vol. 9, No. 3: 185-211, 1990.

Sarrafzadeh, A., Fan, C., Dadgostar, F., Alexander, S. and Messom, C., "Frown gives game away: Affect sensitive tutoring systems for elementary mathematics", *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics,* The Hague, The Netherlands, pp.13-18, October 10-13, 2004.

Sarrafzadeh, A., GholamHosseini, H., Fan, C., and Overmyer S. P., "Using Machine Intelligence to Estimate Human Cognitive and Emotional States from Non-Verbal Signals", *IEEE International Conference on Advanced Learning Technologies* (ICALT), Athens, Greece, pp. 336-337, July 2003.

Sarrafzadeh, A., Messom, C., Overmyer, S., Mills, H., Fan, H., Bigdeli, A. and Biglari-Abhari, M., "The Future of Computer-assisted Management Education and Development", *International Journal of Management Literature,* Vol. 2, No. 4: 214-29, 2002.

Schwarz, N. and Clore, G.L., "Feelings and phenomenal experiences", In E.T. Higgins & A. Kruglanski (Eds.), *Social psychology: Handbook of basic principles*, Guilford Press, New York, USA. pp. 433-465, 1996.

Sirakaya, E. and Sonmez, S., "Gender Images in State Tourism Brouchures: An Overlooked Area in Socially Responsible Tourism Marketing", *Journal of Travel Research*, Vol. 38 (May): 353-362, 2000.

Srikumar, K. and Bhasker, B., "Personalized product selection in Internet business", *Journal of Electronic Commerce Research*, Vol. 5, No. 4: 216-27, 2004.

Stevens, H.P., "Matching Sales Skills to Customer Needs", *Management Review*, Vol. 78, No. 6: 45-47.

Turk, M. A. and Pentland, A. P., "Eigenfaces for recognition", *Journal of Cognitive Neuroscience,* Vol. 3, No. 1: 71-86, 1991.

Vapnik, V.N., *The Nature of Statistical Learning Theory,* Springer-Verlag New York, Inc., New York, 2000.

Williams, D. and Shah, M., "A Fast Algorithm for Active Contours and Curvature Estimation", *CVGIP: Image Understanding,* Vol. 55, No. 1: 14-26, 1992.

Yuasa, M., Yasumura, Y. and Nitta, K., "A Negotiation Support Tool Using Emotional Factors", *IEEE*, Vol. 5: 2906-2911, 2001.