MEASURING CONSUMER INTEREST BY USING LOCAL BINARY PATTERNS

K.Sridhar Reddy*

P.Venkateswara Rao*

Dr. C.Naga Raju**

Abstract—

Using computers to analyze human faces has been an area of recent interest in computer science and psychology. 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. This paper describes an intelligent Approach for finding the costumer interest by analyzing the facial expressions of the customer and performs the required action. Upon entering the shop, a customer has his features scanned and analyzed by the computer and the customer is categorized as a browser, future customer, probable customer or buyer. This feature would also tell the sales personnel whether or not the customer requires or desires assistance in the first place. This, for the customer, can mean being directed to products they have been recognized to be more interested in, resulting in savings in time. In this paper we use Local Binary Patterns (LBP) for face recognition. LBP is a non-parametric kernel which summarizes the local spacial structure of an image and it is a invariant to monotonic gray-scale transformations; hence the LBP representation may be less sensitive to changes in illumination. 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, Face Recognition, facial expressions, Local Binary Patterns non-parametric kernel;

* Assistant Professor, Lakireddy Bali Reddy College of Engineering, Mylavaram-India.

** Professor&Principal, Pace Engineering College, Ongole-India.

I. 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 [1] through facial expressions, gestures, or vocal inflections [2]. 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 [3]. 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. People visiting a store can be broadly classified into four categories [4][5]. 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.

II. LOCAL BINARY PATTERNS

In this section, we introduce the original LBP operator as well as several extensions (multi-scale LBP, uniform LBP, improved LBP) and variants (Extended LBP and Census Transforms).

A. *The Original LBP*

The Local Binary Pattern (LBP) operator is a non-parametric *3x3* kernel which summarizes the local spacial structure of an image. At a given pixel position (x_c, y_c), LBP is defined as an ordered set of binary comparisons of pixel intensities between the center pixel and its eight surrounding pixels

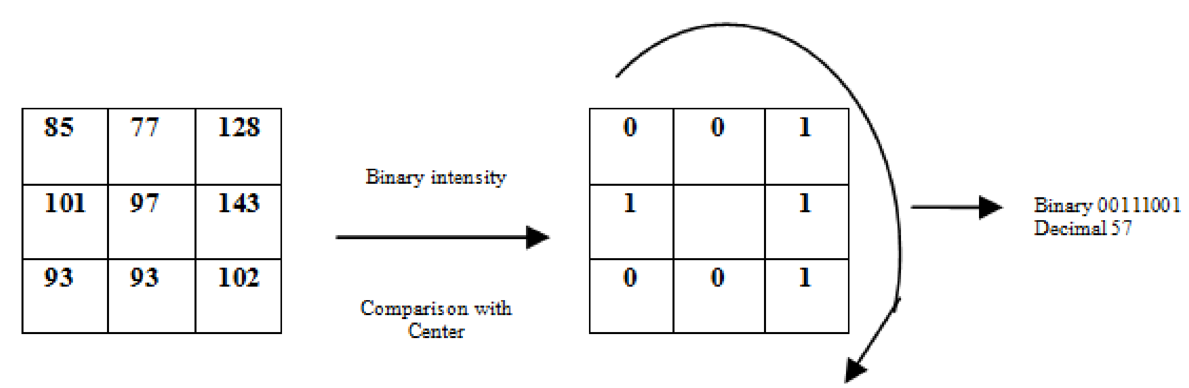

Fig: 1 calculating the original LBP code

The Decimal form of the resulting 8-bit word (LBP code) can be expressed as follows:

$$LBP(x_c, y_c) = \sum_{n=0}^{7} s(i_n - i_c)2^n \quad (1)$$

Where i_c corresponds to the grey value of the center pixel (xc, yc), in to the grey values of the 8 surrounding pixels, and function *s(x)* is defined as:

$$s(x) = \begin{cases} 1 & \text{if} \quad x \geq 0 \\ 0 & \text{if} \quad x < 0 \end{cases} \quad (2)$$

Note that each bit of the LBP code has the same significance level and that two successive bit values may have a totally different meaning. Actually, The LBP code may be interpreted as a kernel structure index. By definition, the LBP operator is unaffected by any monotonic gray-scale transformation which preserves the pixel intensity order in a local neighbourhood.

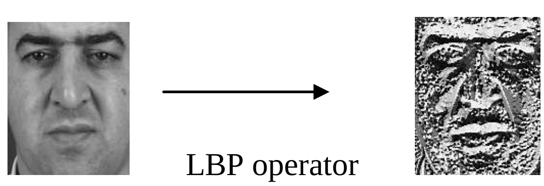

Fig: 2 Original image (left) processed by the LBP operator (right).

Due to its texture discriminative property and its very low computational cost, LBP is becoming very popular in pattern recognition.

B. The Multi-Scale LBP

In the later version original LBP operator extends to a circular neighbourhood of different radius size. Their $LBP_{P,R}$ notation refers to P equally spaced pixels on a circle of radius R. For instance, the $LBP_{8,2}$ operator is illustrated in Fig:3

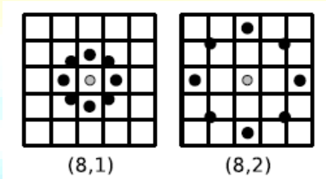

Fig: 3 Examples of extended LBP operators

C. The Uniform LBP

Uniform patterns, contain at most two bitwise 0 to 1 or 1 to 0 transitions (circular binary code). 11111111, 00000110 or 10000111 are for instance uniform patterns. They mainly represent primitive micro-features such as lines, edges, corners $LBP_{P,R}^{u2}$ denotes the extended LBP operator (*u2* for only uniform patterns, labelling all remaining patterns with a single label).

D. The Improved LBP

Recently, new extensions of LBP have appeared. For instance [8], remarked that LBP features miss the local structure under some certain circumstance, and thus they introduced the *Improved* Local Binary Pattern (ILBP). The main difference between ILBP and LBP lies in the comparison of all the pixels (including the center pixel) with the mean of all the pixels in the kernel. The decimal form of the resulting 9-bit word (ILBP code) can be expressed as follows:

$$ILBP(x_c, y_c) = \sum_{n=0}^{8} s(i_n - i_m)2^n \quad (3)$$

Where i_m corresponds to the mean grey value of all the pixels, and function $s(x)$ is defined as in Equation 2.

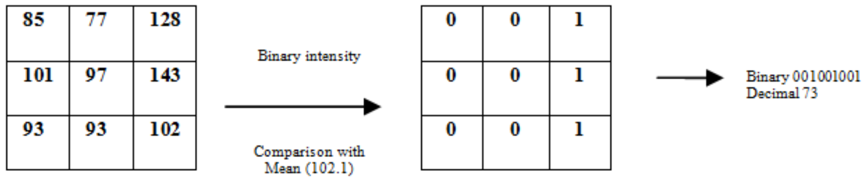

Fig: 4 calculating the ILBP code

E. The Extended LBP

LBP can only reflect the first derivation information of images, but could not represent the velocity of local variation [8]. To solve this problem, they proposed an *Extended* version of Local Binary Patterns (ELBP) that encodes the gradient magnitude image in addition to the original image. For that purpose, they simply applied kernels $LBP_{8,1}^{u2}$, $LBP_{8,2}^{u2}$ and $LBP_{8,3}^{u2}$ both to the original image and the gradient image. As a consequence, this method can't be considered as an extension of the LBP operator.

III. FACE RECOGNITION SYSTEMS USING LOCAL BINARY PATTERNS

The individual sample image is divided into R small non-overlapping blocks (or regions) of same size [10]. Histograms of LBP codes H^r, with $r \in \{1,2,....,R\}$ are calculated over each block and then concatenated into a single histogram representing the face image. A block histogram can be defined as:

$$H^r(i) = \sum_{x,y \in block\, r} I(f(x,y)=i), i=1,....,N$$

(4)

Where N is the number of bins (number of different labels produced by the LBP operator), *f(x, y)* the LBP label at pixel *(x, y)* and *I* the indicator function.

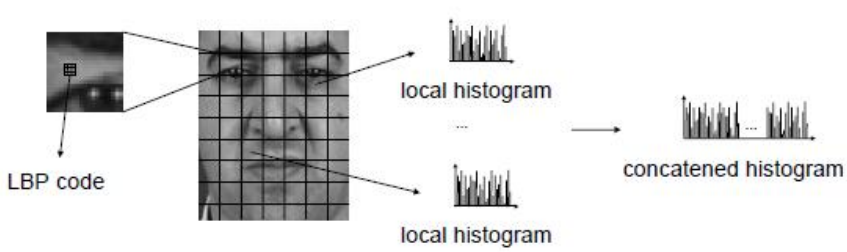

Fig: 5 LBP description of the face

This model contains information on three different levels (Figure 5): (1) LBP code labels for the local histograms (pixel level), (2) local histograms (region level) and (3) a concatenated histogram which builds a global description of the face image (image level). Because some regions are supposed to contain more information (such as eyes), this system propose an empirical method to assign weights to each region. For classification, a nearest-neighbour classifier is used with Chi-square (χ^2) dissimilarity measure, defined as follows:

$$\chi^2(S,M) = \sum_{r,i} \frac{(S^r(i) - M^r(i))^2}{S^r(i) + M^r(i)} \quad (5)$$

Where S and M correspond to the sample and the model histograms. However, the algorithm reported best results with $LBP_{8,2}^{u2}$

A. *Facial Feature Location*

We used regions of interest to divide the face into three small regions as shown in Fig 6. 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. 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.

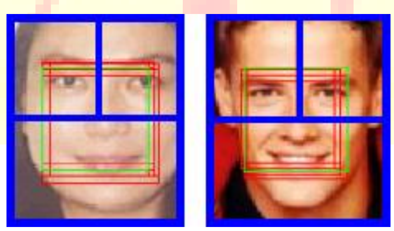

Fig 6: Regions of Interest

B. *Finding the Eye Parameters*

For each possible eye location we start with a minimum possible pupil radius and iteratively apply equation (6) from that location. The aim is to find an *x, y* pair and a pupil radius that maximises

IJESM Volume 2, Issue 2 ISSN: 2320-0294

τ (a ratio of contrast between circles of differing radii). If the pupil's radius grows too large or the optimal eye x, y location moves too far from the original possible eye start point then the current possible eye is rejected as a real eye. If the translations convolve with the centre point of another possible eye point then that other possible eye is removed as a possible eye, but the current possible eye is allowed to translate further before being invalidated.

$$\tau = \frac{\left\{ \dfrac{sum(\underline{x},\underline{y},\underline{r}+1) - sum(\underline{x},\underline{y},\underline{r})}{\underline{r}+1} \right\}}{\left\{ \dfrac{sum(\underline{x},\underline{y},\underline{r}) - sum(\underline{x},\underline{y},\underline{r}-1)}{\underline{r}} \right\}} \qquad (6)$$

Where sum(.) simply sums the pixels in the circle within the input image defined by the arguments. Once τ has been maximised for all of the remaining possible eyes, a repeat of the above algorithm checks that all possible eye pairs have similar radius values and that the inter eye distances are valid compared to the pupil radii etc. The eye pair with the highest τ sum is classified as the real eye pair, with the eye centres and pupil radii recorded. Next we find the eye outers and inners. We normalise only the local area around each eye using (1) and (2). Then a Sobel operator is convolved on this small space to find the eye outline. Blob analysis of the outline structure quickly reveals the extreme top and bottom values (the top and bottom of the eye) and the extreme left and right values.

Fig: 7 Eye Samples

Fig:8 Finding the Eye Parameters. a) Normalise locally,

b) Use edge detection and blob analysis

C. *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.

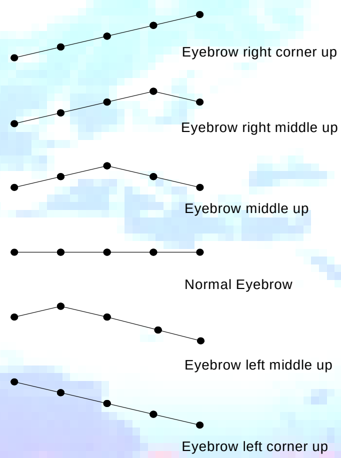

Fig 9: Basic Shapes of Eyebrow

D. *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.

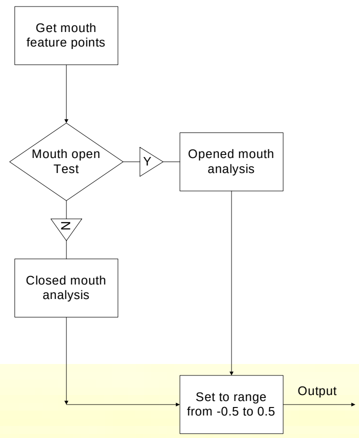

Fig 10: Mouth Analysis Procedures

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’.

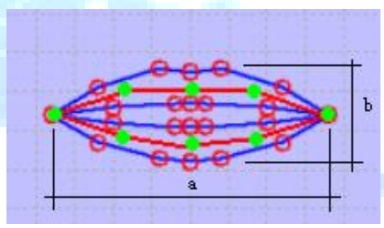

Fig 11: Mouth Model

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. Fig 12 shows these shapes. The output is in the range 0.0 to 1.0.

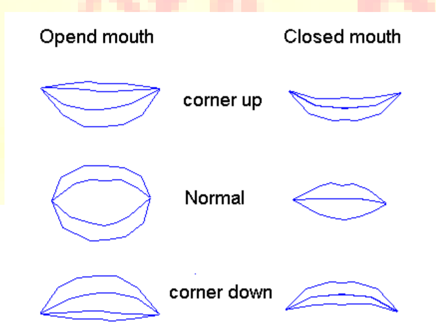

Fig 12: Basic Shapes of the Mouth

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.

E. Related Work

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. The longest bar is the best value for the classifier. Images captured via a webcam are analyzed

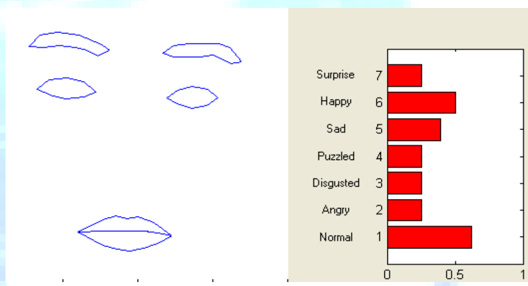

Fig 13: Normal Expression Detected by the Facial Expression Analysis Software

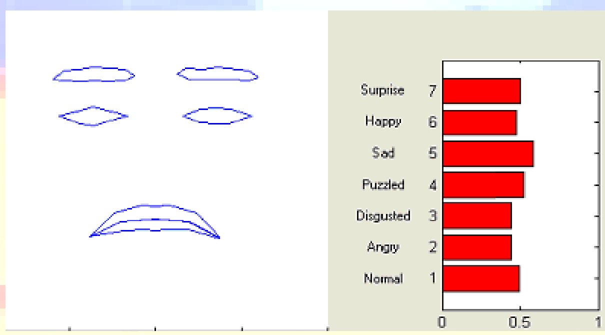

Fig 14: Sad Expression

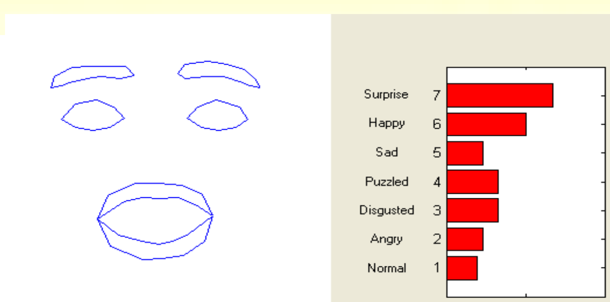

Fig 15: Surprised Expression

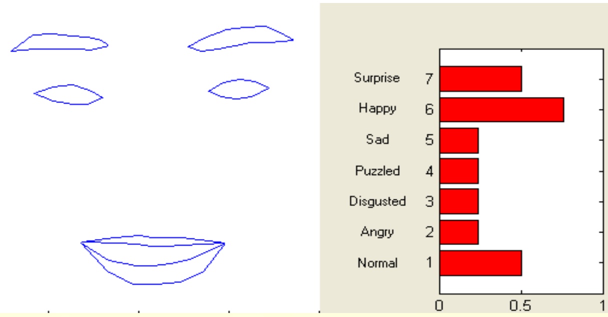

Fig 16: Happy Expression

IV. 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. 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. 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.

REFERENCES

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

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

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

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

[5] 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.

[6] T.Ojala and M.Pietikainen and D.Harwood, "A comparative study of texture measures with classification based on feature distributions",*Pattern Recognition*, Volume 29, 51–59, 1996.

[7] T.Ojala and M.Pietikainen and T.Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with loval binary patterns", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 24, 971–987, 2002.

[8] H. Jin, Q. Liu, H. Lu and X. Tong, "Face detection using improved LBP under bayesian framework", *International Conference on Image and Graphics*, Hong Kong, China. 306–309, 2004.

[9] X. Huang, S. Li, and Y. Wang, "Shape localization based on statistical method using extended local binary pattern", *International Conference on Image and Graphics*, Hong Kong, China, 184–187, 2004.

[10] T. Ahonen, A. Hadid and M. Pietik¨ainen, "Face recognition with local binary patterns", *European Conference on Computer Vision,* Prague, 469–481, 2004.

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

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

[13] 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.

[14] 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.

[15] 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.

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