An Artifical Neural Network for Classification of Color Images

Andrés B. Jaramillo
Pontificia Universidad Javeriana, Engineering Department, Cali, Columbia

Kazuo Yamaba
Mechanical Engineering Laboratory, Robotics Department, Ibaraki, Japan

Abstract

An *Artificial Neural Network* (ANN) with a raw image *Fuzzy* pre-processing mechanism system for static colored pattern classification is presented. A computational three layered feed-forward network utilizing a non-linear supervised learning paradigm is trained on fuzzily processed chromatic and achromatic pattern values. During training, center and bandwidth parameters for the tunning of antecedents in *Fuzzy* rules, corresponding to perceived opponent color categories, are reinforced. These tunned rules are subsequently used to pre-process raw bit test patterns before automatic ANN categorization. By adjusting the opponent primary pairs using the proposed approximate reasoning methodology in conjunction with the ANN, partial human-like visual perception characteristics (primarily color constancy, shape constancy and limited size constancy) are achieved. A particular test bed application has been chosen to demonstrate the usefulness of this system in industrial environments, namely, an automatic visual inspection machine for mounted SMT (Surface Mount Technology) PCB's (Printed Circuit Boards). In this particular application grey-scale inspection proved ineffective due to similar tone scale values of PCBs and some miniature components. Part existence, orientation and correct terminal soldering inspection and classification are being performed under real-time, and real environmental constraints with high hit rates, and low system training trials.

Introduction

Color Perception

The human eye registers as light stimuli of different wavelengths, within around 400 and 700 *nm*, to produce diferent color sensations. Colored light does not exist but it is perceived as visible radiation of different wavelengths. So color perception partially depends upon the wavelength of the light received.

Hue, can be defined to be the psychological dimension that most clearly corresponds to variations in wavelength The purest color one could get would correspond to a monochromatic or spectral hue. As we add other wavelengths, or white light, the color appears to *saturate* or "wash out".

Another important factor, specifically related to the intuitive quantification of color categories, is that of color mixtures, namely *additive* and *substractive color mixtures*. Pigments work by substracting or absorbing wavelengths of light, so mixtures of pigments are called substractive color mixtures. Light, on the other hand behaves as color additive mixtures, which are easier to conceptualize.

Helmholtz (1821-1894) and Maxwell (1831-1879) carried out a set of color matching experiments and reported that by combining an appropiate set of three monochromatic light sources in appropiate amounts, they could match any other hue. These three wavelengths (small, medium, and large within our visible spectrum) were to be known as primaries, and relate directly to the way in which colors are perceived by the retinal cones.

In this proposed system, fuzzy rules are defined that will produce based on three primary color opponent processes a subjectively perceived color. Eventhough we as humans can discriminate, on average, approximately 200 different hues,[1] an appropiate scheme to account for wavelength variations has been implemented using fuzzy reasoning on perceived colors. A mixed approach to specify how a particular color sensation might arise, based on Ewald Hering's Opponent Process Theory [2] and Thomas Young's Trichromatic Color Theory (1802), has been adopted.

In short, Hering's theory states that instead of three primary subjective colors, we have four (including yellow), and that these are arranged in opposing pairs. A third opposing pair was suggested to account for brightness perception. Physiological evidence has supported the fact that neural responses in the eye are subject to excitatory and inhibitory influences caused by interaction between neighboring units. The opposing pairs are given by Red-Green, Blue-Yellow, and Black-White. They are defined as opponent processes because we cannot perceive a redish-green color, or a blueish-yellow (using improper language).

Other schemes have been infered in order to attack specific problems in pattern recognition like color constancy. Chung et al.[3] implement the hypothetical model introduced by Jameson and Hurvich,[4] presenting opponent Lateral Geniculate Nucleus (LGN) units in an ANN configuration. They suggest a method for realizing color constancy by comparing the inputs from the opponent LGN units in order to identify the spectral reflectance of visual objects. Another well known study is that of Land in his "Retinex Theory",[5] and others that have emerged based on Land's work. A good number of these studies have used Hering's and Young's theories as basis for their work, showing for example high orthogonality of spectral response properties between opponent color processes.[6]

The Constancies

Objects seen by our visual sense remain in a permanent position despite the fact that their images fluctuate within our eye (even when staring directly into an object, human observers execute between three to four eye movements per second). You feel as though you draw closer to objects of a fixed size instead of feeling changes in the size of the object itself, and objects seem to be the same color and brightness regardless of the intensity of the light falling upon it.

Since the physical world does not change as the light, or our direction of gaze, or our distance from a target changes, we seem to construct a corresponding stable world of consciousness eventhough sensory properties vary. The way in which this stability is created and maintained revolves around a set of visual phenomena known as the *constancies*. There are constancies related to visual objects, such as their size and shape;[7] constancies related to qualities of those visual objects, such as their brightness (perceived phenomena that is directly related to luminance and reflectance) and hue; and constancies that deal with relationships between visual objects, such as their position or orientation.

Needless to say our artificial perception systems lack most of these abilities, hence the enormous difficulty in digital image understanding of real objects.

In general terms these constancies are intrinsic properties of our visual perception and come about due to numerous and complex interactions between our senses and our psychological behavior. These reasons lead us into the study of the mechanisms that lay behind the constancies, and that must be understood before we can completely emulate them.

In this paper an attempt is made to emulate such properties, by assigning fuzzy membership functions to respresent subjective primary color categories in accordance to the opponent process theory, thus "extending" the boundaries of known (in system's memory) objects in order to provide some permitted range of comparison between observed and learned data instead of performing direct raw pattern matching. Objects within an image can be slanted, bigger, and even color-hue-different to some extent, and still be recognized.

Fuzzy Logic

Most of the existent traditional tools for formal modelling, reasoning, and computing are crisp, deterministic and precise in characte.r[8] That is, binary true-false type rather than approximate more-or-less type. Precision assumes that the parameters of a model respresent exactly either our perception of the phenomenon modelled or the features of the real system that has been modelled (i.e. it contains no ambiguities). It is important to note that:

> Real situations are very often not crisp and deterministic and they cannot be described precisely by conventional models. The discrepancies between human perceived sensory information and artificial transducers [9] are not deterministic in nature, hence no exact mathematical model can relate both domains.
>
> The complete description of a real system often would comprise far more information than a human being could ever recognize simultaneously, process and understand.

As L. A. Zadeh puts it[10]: "As the complexity of a system increases, our ability to make precise and yet significant statements about our behavior diminishes until a threshold is reached behond which precision and significance (or relevance) becomes almost mutually exclusive characteristics".

In real situations, uncertainty and vagueness can be accounted for in a number of ways. Uncertainty can come from the lack of information about future states of a system, and it can be handled appropiately (in most cases) by probabilistic theory and statistics, or it can also be found on the degree of truth of statements and therefore bases on logic. On both of these types of probabilistic approaches it is assummed, however, that the events (elements of sets) or statements are well defined (*stochastic uncertainty).*

The vagueness concerning the description of the semantic meaning of events, statements themselves, is what is usually called *fuzziness*.

It is clear that fuzziness can be found in most of our daily communication using natural languages as well as in our ways of thinking and behaving.

In order to emulate human behavior, it is necessary to represent intuition and uncertainty in the know-how, subjective perception of surrounding data, and skills. These cannot be included in probability but fuzziness. Probability is based on the estimation of the degree at which the event will easily occur, before the occurrence and it is characterized by a probability density function. Thus the probability is meaningless after the occurrence. On the other hand, intuition such as "she is beautiful", "it is too hot", "it is dark red", etc., includes another type of ambiguity which cannot be clear after the occurrence and depends upon the person. This can be included in Fuzzy sets (or fuzzy subsets) and characterized by a *membership function*.. Fuzzy inference, or approximate reasoning, is a simulation of decision making of human experts.[11]

Linguistic terms, whose definitions are not so clear, are used commonly for easy and effective communication. Intuitive and vague terms that can be understood by the *common sense* assigned by communicating persons, are very easy to select for practical use, although they include uncertainties. Linguistic terms can be defined by a *membership function*, which indicates a grade of membership of each element in a fuzzy linguistic term of interest. These grades of belonging are given by intuition or common sense, so that the shape of a given membership function varies a little from person to person. Fuzzy linguistic terms can be defined by membership functions exhibiting a continuos curve changing between 0 and 1, where 0 corresponds to *no belonging* and 1 *completely belongs*. The transition region represents a fuzzy boundary of the specific term. Fuzzy sets are defined by *labels* (e.g. a little, a lot, and so on) and membership functions.

In mathematical terms, X being a collection of objects denoted generically by x then the fuzzy set $\tilde{A}$ in X is a set of ordered pairs:

$$\tilde{A} = \left\{ \left(x, \mu_{\tilde{A}}(x) \right) \middle| x \varepsilon X \right\} \tag{1}$$

where $\mu_{\tilde{A}}(x)$ is called the membership grade of x in $\tilde{A}$ which maps x to membership space M. When M contains only points 0 and 1, $\tilde{A}$ is said to be non-fuzzy. Fuzzy linguistic terms, elements of which are ordered, are fuzzy intervals. Elements giving grades of membership to be at

0.5 are *crossover points (see Fig. 1)*, and the interval between crossover points is the *bandwidth* of the fuzzy linguistic term. An interval on the horizontal axis where the grades of membership are not zero is called *support*.

Knowledge acquisition and representation

Two mayor fields of study related to information processing are *knowledge acquisition* and *inference*. Knowledge is usually represented by the relation from causes to effects in the form of **IF-THEN** rules.

Knowledge acquisition in human beings is not necessarily achieved by exact matching between input data and antecedents in the knowledge base. Human beings can gain a much smaller amount of useful knowledge from numerous experiences by what has been defined as a *summarization process*. Summarization is to cut off the less important portions from raw information, to emphasize more important points and to extract the essence of the information. Summarization process converges many similar pieces of information obtained from experiences into one simple piece of information which includes a very important essence. This in turn also allows us to store a smaller amount of know-how efficiently. It is also a significant fact that know-how obtained by summarization is usually represented by fuzzy linguistic terms. Otherwise, know-how will be the expression of only one experience and reduction of it cannot be guaranteed.

Thus knowledge acquisition is based upon *summarization* and *fuzzification*, and it is precisely the way data acquisition and learning are presented to the system proposed. Both, *summarization and fuzzification*, are implicitly considered within this combined fuzzy inference, and ANN system.

Why Fuzzy Pre-processing?

The approach developed centers on the functional behavior of the human vision, and not on the hardware/software replication (brute force) of human visual "transducers" and brain's "processor architecture".

The utility of data for more complex downstream processing tasks such as clustering and classifier design can be improved for feature analysis during fuzzy preprocessing of the "raw" data[12].

On the other hand, since human-like behavior is expected of such a system (positive type behavior of course!) we need to find ways of expressing information in a similar manner. Fuzzy theory is one alternative that can provide us with good tools to tackle the problem. It not only reduces quantitative representation complexity (quantification and qualification of knowledge), but also permits faster modelling of the studied or processed system characteristics. In general terms:

- Faster and simpler knowledge representation
- Can be tuned to obtain alternate behaviours even based on intuition (i.e. flexibility)
- It can accept compound information obtained from incomplete input data and still produce reasonable conclusions.
- Good ability to interpolate between input-output relations

Let us assume that each primary color can be defined by three general hue linguistic categories, namely Dark (**D**), Normal (**N**), and Light (**L**). Other primaries can be selected intuitively from the CIE color charts, but in this case the hardware utilized works directly with near red, green, and blue frequencies. They are obtained via digitalization from an NTSC standard frame signal, individually decoded into the four components required (3 chromatic and 1 achromatic).

Figure 1 shows the Fuzzy characteristic function, or membership function for the antecedent of the IF-THEN clauses. They represent the fuzzy terms employed to describe the perceived hue values of the primary colors and luminance (R,G,B,L).

The dotted lines represent the maximum tolerable bandwidth of the corresponding label, while the normal lines correspond to the default settings when the fuzzy-neural system begins the tuning and training of fuzzy rules and neural interconnections. The vertical axis, labelled μ, defines the degree of membership ranging from 0 to 1, and the horizontal axis the hue value. In the actual implementation the horizontal axis is defined in a normalized 0-255 range, in part due to the image capture board A/D conversion word size (8 bits).

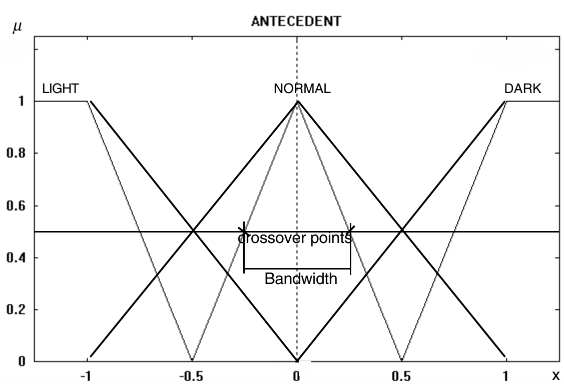

Figure 1. Membership functions primary color hues

Table 1.

	Opponent2		
Opponent1	***D***	***N***	***L***
D	**AZ**	**MN**	**N**
N	**MP**	**AZ**	**MN**
L	**P**	**MP**	**AZ**

Where;

AZ: Around Zero
P: Positive
N: Negative
MN: Medium Negative
MP: Medium Positive

These correspond to the singleton linguistic labels of the consequent as defined in the cross-reference table 1. Consider the following fuzzy inference for the red-green opponent pair:

Rule No. 1:	**IF** *RED=D* **and** *GREEN=D* **THEN** *opponent(R,G)=AZ*
Rule No. 2:	**IF** *RED=D* **and** *GREEN=* N **THEN** *opponent(R,G)=MP*
Rule No. 3:	**IF** *RED=D* **and** *GREEN=L* **THEN** *opponent(R,G)=P*
Rule No. 4:	**IF** *RED= N* **and** *GREEN=D* **THEN** *opponent(R,G)=MN*
Rule No. 5:	**IF** *RED=N* **and** *GREEN=N* **THEN** *opponent(R,G)=AZ*
Rule No. 6:	**IF** *RED=N* **and** *GREEN=L* **THEN** *opponent(R,G)=MP*
Rule No. 7:	**IF** *RED=L* **and** *GREEN=D* **THEN** *opponent(R,G)=N*
Rule No. 8:	**IF** *RED=L* **and** *GREEN=N* **THEN** *opponent(R,G)=MN*
Rule No. 9:	**IF** *RED=L* **and** *GREEN=L* **THEN** *opponent(R,G)=AZ*

Accordingly, each of the remaining two opponent processes (Blue-Yellow, and Black-White) has 9 fuzzy rules that define its input-output relationship in linguistic terms. These IF-THEN clauses have a conjunction of two fuzzy linguistic terms in the antecedent and a singleton consequent (Figure 2).

The mechanism of fuzzy inference when a fact is given as a fuzzy value can be obtained by a matching grade, taking a MAX of MIN of a fact and a fuzzy variable in an antecedent. The matching is a soft matching.[11] In this case we have two variables in each antecedent, and the minimum of their matching grades will be the matching grade of the antecedent. Individual results for each antecedent are obtained by truncating a consequent with the matching grade. Furthermore, the final conclusion can be obtained by ORing these individual results. In those cases were a crisp output is required by the system (most practical applications) the Center of Gravity, given by formula (2), can be used as a method to achieve *defuzzification*.

$$C.G = \frac{\int i\mu(i)di}{\int \mu(i)di} \tag{2}$$

for continuous expressions of a membership function, where (i) is a grade of ith element.

Since the application in hand deals with deterministic values for the inputs and outputs. That is, raw pixel values are given as inputs so a matching grade is directly obtained from the grade of a membership function in an antecedent. And, the final conclusion is obtained by the Center of Gravity method described by:

$$C.G = \frac{\sum_{i=1}^{n} i\mu_i}{\sum_{i=1}^{n} \mu_i} \tag{3}$$

Figure 2 shows, singleton labels in solid lines and triangular functions (not used) in dotted lines. Singletons where preferred for the consequent since it gives rise to a linear interpolation between antecedents and consequent[17]. f(op1,op2) is the function relating the two opponent pairs selected.

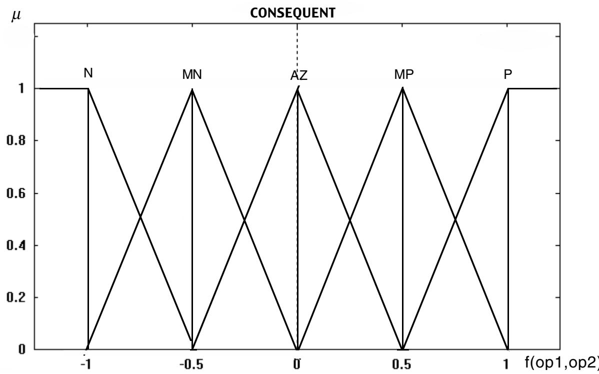

Figure 2. Consequent given by opponent processes

Artificial Neural Networks

Why NeuroComputing?

When it becomes expensive, or even impossible in some cases, to determine a systems exact mathematical model (such is the case with humans visual system), we cannot easily define it's input-output relational behavior. ANNs provide a mean to "learn" and create a systems characteristic model based on known, or even unknown (depending on the learning paradigms used), input-output response pairs. This means, we can treat our eye-brain like a black box with relatively known perceived behavior to seen data. In other words we can teach a network how to behave when confronted with new "unknown" input data (Physiological, and psychological). All this means is that ANNs facilitate interpolation between input-output pairs.

Furthermore, any function $fe\colon \Re_p \rightarrow \Re_q$ where p >= q is a feature extractor when applied to X. The new features are the image of X under fe. Hence, Fuzzy Inferencing and ANN can be considered to be appropiate feature extractors.

Figure 3 shows a typical processing element configuration in ANN. Figure 4 shows a typical feedforward configuration of an ANN, where f is the transfer function (usually a sigmoid function) applied to the summation of aggregate node weight-value products in the network. A learning paradigm is applied in order to adjust interconnection weights between neurons.

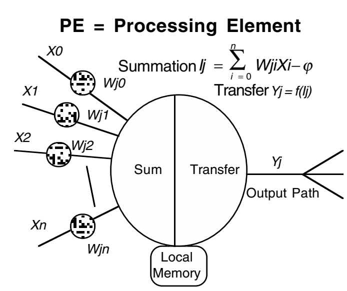

Figure 3. Building blocks of a simple ANN

Learning Algorithm

Most of the best-known learning paradigms[13,14,15,16] have been implemented and tested with the combined fuzzy-ANN system. The best results have been obtained using a variation of the *Backpropagation*[13] algorithm described subsequently.

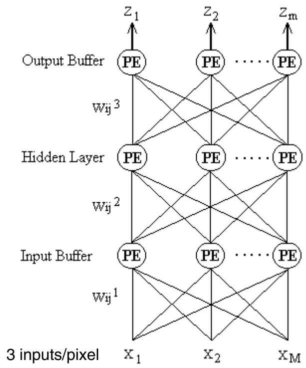

Figure 4. A typical feedforward ANN

Backpropagation is characterized by inputting data into the bottom of a multilayer network, then propagating the information through the network in a feedforward manner. When the output nodes are computed, their output is compared to the desired classification of the input data. The error is then used to compute a correction for the interconnection weights between network nodes. The correction is just a gradient descent down an error surface. The error is backpropagated through the network, correcting the weights. It is not a "natural" algorithm since human networks cannot backpropagate error and adjust this way but it gives desirable results at low computational costs in ANN.

Because this algorithm uses a fixed learning constant and computes the gradient for each individual input vector as opposed to computing the gradient for the entire set of training data, the error might not be reduced by successive iterations. For this reason backpropagation cannot be proven to converge. As in all gradient descent algorithms it may get caught in a local minima solution so techniques like simulated annealing, and random presentation of input classes have been used to improve its performance.

Learning is performed according to the following equation:

$$\omega_{ij}^{+} = \omega_{ij}^{-} + \eta \delta_j \chi_i + \alpha(\omega_{ij}^{-} - \omega_{ij}^{--}) + m_{ij} \qquad (4)$$

where ω_{ij} is the weight from node i to node j in the next layer, x_i is the output of node i, and δ_j is the error associated with node j. η and α are learning rates. ω_{ij}^{+} is the new weight value and ω_{ij}^{-} is the old weight value. ω_{ij}^{--} is the value of the weight before the last update. Thresholds are adapted similarly where x_i is replaced by +1 if the threshold is added to the weighted sum and -1 if it is subtracted. Finally, a term m_{ij} is added to accelerate convergence since it actuates as a momentum term for the normal gradient descent backpropagation learning algorithm.

The δj are defined as follows:

$$\delta_j = \begin{cases} y_j(1-y_j)(d_j-y_j) & \text{output node j} \quad (5) \\ x_j(1-x_j)\sum_k \delta_k \omega_{jk} & \text{hidden node j} \quad (6) \end{cases}$$

where dj is the desired output for the output node j and yj is the actual output. For the hidden nodes the δj are the errors for the layers above.

The Implementation

The Fuzzy-Neuro System has been implemented in software under an IBM PC compatible platform. It has been designed as a task driven program that allows for the application of ANN solutions to almost any appropriate problem.

Learning and test data can be supplied manually or automatically. During execution, the network can be asked for current and past states. Teaching and operating parameters can be modified at any time in order to change the behavior of the system. It also includes user controlled parameters like noise excitation. The presentation of inputs can be done sequentially or on a random basis, with simulated annealing implemented to prevent the local minima problem.

It provides an optimum mechanism for solving "unknown" input-output system relations. Solutions for logical and non-logical systems are obtained with relative ease.

A particular module developed for image recognition permits the ANN to learn fuzzily pre-processed colored image patterns captured by a CCD camera (as well as MOS cameras or other type of visual sensor) and then test different "new" patterns for their categorization. By fuzzily pre-processing raw bit pixel information based on subjective human color categories, partial shape, color, and size constancy have been achieved. A compression of data is also managed that allows NTSC decoded signals (in their primary chromatic and achromatic components) to be treated at 3/4 of their actual storage size.

During training of the ANN, the fuzzy rules corresponding to the opponent color processes are tuned based on the pixel values of each primary color and luminance. Label and bandwidth values of the fuzzy rules are adjusted according to the training data set, based on the tought class in the supervized ANN i.e. samples from class x only affect class x's fuzzy rules. In this way, further separation between class patterns can be achieved permitting faster convergence to an optimum solution space while training the system. The selected shapes for the fuzzy rules have been chosen to be the S-function and the Z-function since the can be easily and rapidly computed, and they can also be combined to form trapezoidal and triangular functions. It has also been decided this way because the system might be completely transported to the hardware domain using a set of Fuzzy chips, Omron FP9000 and FP9001, that adjust to these specs.[17]

The fuzzy pre-processing has been implemented in our proprietary PROLOG-like interpreted language (Edinburgh format) developed under Borland's Turbo Prolog (V2.0) compiler. This language can be appropriately designed to use subjective linguistic terms and produce deterministic results.

The ANN has been implemented in Borland's Turbo C++ (V3.1). Direct memory exchange of data, or by file handling, between the two environments takes place within the same package (i.e. data produced by the fuzzy pre-processing can be directly accessed by the ANN via memory, or via files). Network topology can be defined by the user and network parameters as well. Embedded learning rules give the user a flexible set of alternatives when training and testing of the network is requested.

The Specific test bed application

The developed Neuro-Fuzzy System has been tested for several toy problems mainly dealing with classification and binary logic inferences. Furthermore, and as mentioned earlier it has been proven effective in a real practical application developed for Kyushu Matsushita Electric Co.

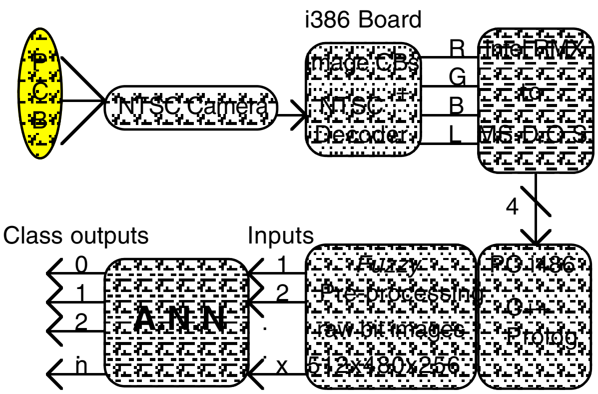

Figure 3. Building blocks of Experimental setup

The application, deals with the inspection of SMT (Surface Mount Technology) PCBs (Printed Circuit Boards) in real time. Figure 3. shows the general setup for processing of images taken from PCBs. With the aid of an NTSC color CCD camera and four (R,G,B,L) fast image capture boards controlled by a i386 (or i486 in future implementations) processor, and an xy positioning mechanism, it has been possible to inspect and classify (good, bad with specific error report) high density mounted PCBs within the tolerable time constraints.

Inspection is carried out in order to detect the following conditions automatically:

- Existence or non-existence of component
- Location and correct placement of component
- Soldering conditions of terminals in component

The following problems have been considered:

- Grey scale images don't provide enough information and can be ambiguous,
- Conventional image processing techniques are too computationally expensive for the real-time constraints involved,
- Miniature components are extremely difficult to recognize, in particular when background color seems to be the same,
- Positioning system must be highly accurate and fast (Direction constancy),
- Shape, color, and size constancy do not hold true in artificial image processing. Cameras and computers do not see what humans see.
- Variations in environmental conditions within in dustrial sites, (especially light sources) affect image recognition.
- Dense PCBs increase the complexity.

Results

The implementation has successfully managed to correct several of these mentioned deficiencies with high hit rates in the classification of different images. Components can be different in color hue, slightly rotated over base axis, or partially viewed (up to 25%, 20%, and 30% respectively). These characteristics can be adjusted to fit the inspection's specifications by manually limiting the fuzzy inferences.

The best result were obtained utilizing a 3 layered fully interconnected feedforward network with 15-20 hidden layer nodes, a tangential or a sigmoid threshold function, and the modified backpropagation learning algorithm. Image size can vary between the different recognition tasks stated in the previuos section. For example for existence determination of 5 *mm* in length transistors mounted over dark background PCB, a normalized image of 8×8 pixels was used (576 inputs to the ANN, and 2 class outputs - component present or not present). The results gave a 100% correct classification rate for under 600 learning trials with an +80% certainty of class selected. Above 1000 training trials the certainty of class selected surpassed 93% and in some cases went up to 98%. The input learning set consisted of raw bit images, with chromatic (3) and achromatic (1) values for each pixel, of PCB sections with components mounted (normally and slanted), components not mounted but soldering paste applied on the face of the PCB, and with PCBs not mounted and with no soldering paste applied.

References

1. Coren, S., Porac, C., Ward, L. M., *Sensation of Perception*, Academic Press, Inc., 1979.
2. Hering, E. *Outlines of a theory of the ligth sense*, 1878. English trans. by L. M. Hurvich and D. Jameson. Harvard University Press, 1964.
3. Chung, C. S., and Han, K. H. "Neural Networks for color vision", Proc. of the Int. Conf. on Fuzzy Logic and Neural Networks, Iizuka, Japan, 1990.
4. Hurvich, L. M., and Jameson, D. "Opponent processes as a model of neural organization", *American Psychologist*, 1974, **29**, 88-102.
5. Land, E. H. "The retinex theory of color vision", *Scientific American*, 218 (**6**), 108-128, 1977.
6. Nakauchi, S., Usui, S. and Miyake, S. "A three layered neural network model which simulates color opponent processing", *Proc. of the Int. Conf. on Fuzzy Logic and Neural Networks*, Iisuka, Japan, 1990.
7. Epstein, W. and Park, J. N. "Shape constancy: functional relationships and theoretical formulations", *Psychological bulletin*, 1964.
8. Zimmerman, H. J., *Fuzzy Set Theory and Applications*, Kluwer-Nijhoff Publishing, 1986.
9. Sagawa, K. and Takeichi, K. "System of Mesopic photometry for evaluating lights in terms of comparative bright ness relationships". *J. Opt. Soc. Am. A*, Vol. **9**, No. 8, 1992.
10. Zadeh, L. A., "Fuzzy Sets", *Information and Control* **8**, 338-353, 1965.
11. Yamakawa, T., "A Fuzzy Logic Controller", *Journal of Biotechnology*, **24**, 1992.

12. Besdek, J. C. "A review of probabilistic, fuzzy, and neural models for PR". *Journal of Intell. and Fuzzy Systems*, Vol. **1** (1), John Wiley & Sons, 1993.
13. Werbos, P. J. "Beyond regression: new tools for prediction and analysis in the behavioral sciences". Doctoral disserta tion, Appl. Math., Harvard Un., 1974.
14. Rosenblatt, F., "The Perceptron: A Probabilistic model for information storage and organization in the brain", *Psychol. Rev.*, **65**, 386-408, 1958.
15. Widrow, B., and Hoff, M. E., "Adaptive Switching Circuits", *1960 IRE WESCON Convention Record*, 96-104, New York, 1960.
16. Hebb, D., *The Organization of Behavior*, Wiley, New York, 1949.
17. Yamakawa, T. "Silicon Implementation for a novel high-speed fuzzy inference engine: mega-flips analog fuzzy processor", *Journ. of Intell. and Fuzzy Systems*, Vol. **1** (1),1993.