Face Detection by Neural Networks Based on Invariant Moments

Romana CAPOR-HROSIK
Faculty of Appl./Business Comp.
University of Dubrovnik
Bran. Dubrovnika 29, Dubrovnik
CROATIA
romanacapor@yahoo.com

Milan TUBA
Faculty of Computer Science
Megatrend University Belgrade
Bul. umetnosti 29, N. Belgrade
SERBIA
tuba@ieee.org

Mirjana VUKOVIC, Milenko PIKULA
Department of Mathematics
University of Sarajevo
Alekse Santica 1, Pale
BOSNIA and HERZEGOVINA
mirvuk48@gmail.com

Abstract: Neural networks beginnings were in the forties, theirs early renaissance experienced in the eighties and today are widely used in pattern recognition, analyzes of medical tests, handwriting, speech, prediction market prices, criminological military research, psychiatric evaluations, data analysis, finding optimal solutions, robot control, weather forecast and many other areas. The great potential of neural networks lies in the possibility of parallel processing of data. Perhaps it will be theirs period of renaissance. This paper introduced an experimental evolution of the effectiveness of utilizing moment invariants as pattern features in human face technology. In this paper we used skin detection and Hu moments for human face localization, also we used neural networks as classifier for this application. The utilized network is a multilayer perceptron (MLP) with one hidden layer. The backpropagation learning is used for its training. Experimental results of neural networks demonstrated successful detection.

Key-words: - Image processing, Face detection, Skin color detection, Hu moments, Neural networks

1 Introduction

Face detection is very active researching field with various applications. Due to technical evolvement numerous of these applications become an irreplaceable part of our daily lives. Mention some, on mobiles we have androids that recognize that user is looking at them so the screen won't turn off, there are modern human-computer interfaces that recognize body postures, body gestures and even facial expressions, etc. Human eye without difficulty can detect another human face unlike computer vision. For it there are many difficulties (mustaches, glasses, beards hide or change same basic features of the face; technical properties of devices,...) and even more challenges (pose variance, imaging conditions, face images can become significantly different under various lighting,...) associated with face localization. Numerous different algorithms have been proposed but the results are still not satisfactory for use in everyday life. . Therefore, face detection is still in its infancy, evolving all the time.

Yang, Kriegman and Ahuja [1] proposed classification for face detection, which is generally accepted. According to them, there are four categories of classification algorithms for face detection: knowledge-based, feature-invariant, template matching based and appearance-based.

Various face detection techniques have been used accomplishing different results [2], [3]. Terrillon et al. [4] presented algorithm for skin color modeling in TSL (tint-saturation-luma) color space that has achieved the best results (in all color spaces for skin color modeling). DeDios and Garcia [5], analyzing YCbCr color space, describe YCgCr color space (use Cg color component instead of Cb component) that accomplish better results than YCbCr. Kukharev and Novosielski [8] built skin color model in YCbCr (people of black skin color have not been investigated). They were defining and comparing thresholds of each Y, Cb and Cr components, what leads to identification of skin/non skin pixels.

In this paper we will describe our algorithm for face detection by neural networks based on invariant moments. First we roughly searched areas of the skin color, and then applying the Hu moments for enhancing its efficiency. Based on the calculated values of Hu moments we built neural networks that classify face from other exposed skin body parts.

The rest of the paper is organized as follows. Section 2 describes color space for skin detection. Section 3 shows description of skin detection algorithm. Section 4 describes briefly Hu moments and their application in our work. Section 5 describes shortly Neural Networks with an emphasis on Multilayer perceptron (MLP). Section 6 discusses the proposed method in detail. Section 7 reports our experimental results. Section 8 concludes the paper.

2 Color space for skin detection

Many models of color spaces have been proposed for skin detection. Some of the most used models of color spaces are YCbCr, HSI, TSL, RGB. There is a proof that if an optimum skin detector is designed for every color space, then their performance will be the same. However, it is generally agreed that there does not exist an one color space which is convenient for detection of skin color in all color images [7].

YCbCr color space is utilized by our algorithm of skin detection. In the name of YCbCr color space Y presents the luminance channel and Cb and Cr are the blue-difference and red-difference chrominance components respectively (Fig. 1). It has emerged as a response to the increasing demands for digital algorithms in handling video information. Except YCbCr, in the family of television transmission color spaces belong YUV and YIQ.

Fig.1. Color image and its Y, Cb and Cr components

Applying the following formulae we get direct conversion from RGB to YCbCb:

$$Y = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B \tag{3}$$

$$Cb = 128 - (0.168736 \cdot R) - (0.331264 \cdot G) + (0.5 \cdot B) \tag{4}$$

$$Cr = 128 + (0.5 \cdot R) - (0.418688 \cdot G) - (0.081312 \cdot B) \tag{5}$$

We had a problem with some converted value of pixels as they came out of their potential range. Therefore it was necessary to return these values within the interval given sizes. We did it in a way that those values that go below zero we put to zero and those that exceeded 255 we set on 255.

3 Skin detection in YCbCr

People have different skin colors in appearance. Numerous studies have showed the main difference lies in intensity rather than color itself. In contrast to RGB, the YCbCr color space is lumaindependent so it is one of the most popular color spaces for skin detection. According to (Hsu et al, 2002), the skin color cluster is more compact in YCbCr than in other color space.

First what we do in our skin detection algorithm is applying color segmentation on an input image (Fig. 2), by using the threshold for each of the components with following inequality

$$69 < Y < 215 \tag{9}$$

$$94 < Cb < 126 \tag{10}$$

$$139 < Cr < 17 \tag{11}$$

Our result is binary image where all pixels are marked either as skin or non-skin (Fig. 3).

Fig.2. Original image

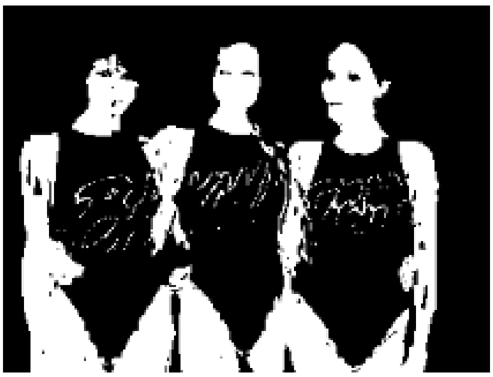

Fig.3. Binary image showing skin areas

4 Moment invariants. Hu moments

Moment invariants were firstly introduced to the pattern recognition field in 1962 by Hu [8], who employed the results of the theory of algebraic invariants and derived his seven famous invariants to rotation of two-dimensional objects.

In pattern recognition, moments and functions of moments have been extensively used as invariant global features of images. An essential feature of pattern analysis is the recognition of objects and characters regardless of their size, position and orientation.

Regular moment invariants are one of the most popular and widely used contour-based shape descriptors, a set derived by Hu (1962) [10]. Later on these geometrical moment invariants have been extended to larger sets by Wong & Siu (1999) and to other forms (Dudani et al 1977;Liao & Pawlak 1998) **Error! Reference source not found.**.

Two-dimensional moments of order $(p+q)$ of digital image $f(x,y)$ of size $M \times N$ are defined as:

$$m_{p,q} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} x^p y^q f(x,y) \tag{12}$$

where $p,q = 0,1,2\cdots$.

The corresponding central moment of order $(p+q)$ is defined as:

$$\mu_{p,q} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)(x-x_{avg})^p (y-y_{avg})^q \tag{13}$$

where $x_{avg} = \dfrac{m_{10}}{m_{00}}$ and $y_{avg} = \dfrac{m_{01}}{m_{00}}$.

The normalized central moments are defined as:

$$\eta_{p,q} = \frac{\mu_{p,q}}{m_{00}^{\frac{p+q}{2}+1}} \tag{14}$$

The first invariants that appeared in literature were invariants to similarity transformation group. They were response to the problem of choosing the crucial properties of objects classification regardless of their position, primarily because of its uncomplicatedness in application. Invariants to translation and scaling are trivial – central and normalized moments themselves solve it. So the only non-trivial problem remains finding rotational invariants. The problem is solved by M.K. Hu who defines the following seven rotational invariants which are computed from central moments through order three.

$$h_1 = \eta_{20} + \eta_{02} \tag{15}$$

$$h_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \tag{16}$$

$$h_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \tag{17}$$

$$h_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \tag{18}$$

$$\begin{aligned} h_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})((\eta_{30} + \eta_{12})^2 \\ &- 3(\eta_{21} + \eta_{03})^2) + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \\ &(3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2) \end{aligned} \tag{19}$$

$$\begin{aligned} h_6 &= (\eta_{20} - \eta_{02})((\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2) \\ &+ 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \end{aligned} \tag{20}$$

$$\begin{aligned} h_7 &= (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})(3(\eta_{30} + \eta_{12})^2 \\ &- (\eta_{21} + \eta_{03})^2) - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03}) \\ &(3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2) \end{aligned} \tag{21}$$

Hu invariants are not demanding for the implementation and therefore they are often used as a basic set of characteristic properties in solving many different pattern recognition problems. Despite numerous disadvantages, range of their application is huge from identification persons, medical applications, objects for robot orientation, recognition of letters and mails, the analysis of satellite images, etc.

Fig.4. One of binary images on which we applied Hu moments

5 Neural networks

There are categories of problems for which cannot be formulated algorithm, problems which require learning to get results, or problems that have exponential time complexity. To solve these kinds of problems we are using neural networks.

At first sight it looks perfect, but on the other side is very complex, because of complex nature of its calculation, so it is very demanding to find fault. Usually because of that problem they are being tested on problems for which there are precise solution algorithms or their solution is known in advance.

Artificial neural networks can be characterized as computational models with particular properties such as the capability to adapt or learn to generalize or to cluster or organize data and which operation is based on parallel processing. However many of the above referred characteristics can be attributed to existing non neural models. The question is to which scope the neural approach demonstrates to be better suited for certain applications than existing models. To this day a reliable answer to this question is not found.

First beginnings of neural computing is usually associated with the 1943rd year and neuro-psychologists Warren McCulloch and Walter Pitts, who had produced the first artificial neuron. Technology available at that time did not allow them to do more in this field. The concept of neural networks was first proposed by Alan Turing in his paper "Intelligent machine "1948th year. In mid-1960th Minsky and Papert in the book "Perceptron" offered evidence that neural networks cannot learn the XOR operation. Thus slowed down the development of this area. Their proof was later disproved, of course. Although their beginnings are in the forties, their development began in the eighties when algorithms become good enough to use. Nowadays, development of the field of research has almost been explosive.

Artificial neural networks have the structure, function and process information similar biological, although the mathematical model of biological network is far more simplified. They are trying to replicate only the simplest elements of this complicated and powerful system. Although they work in a primitive manner, they are very successful in solving many problems.

In this paper Multilayer perceptron (MLP) is used as classifier in face detection system. MLPs are the most popular type of artificial neural networks (ANN). They are feed forward networks of simple processing elements or neurons. Because of their capacity to learn complex non-linear input–output relationships and ability to generalize any given data they have been successfully applied in various pattern recognition problems. The key power provided by such networks is that they admit fairly simple algorithms where the form of the nonlinearity can be learned from the training data.

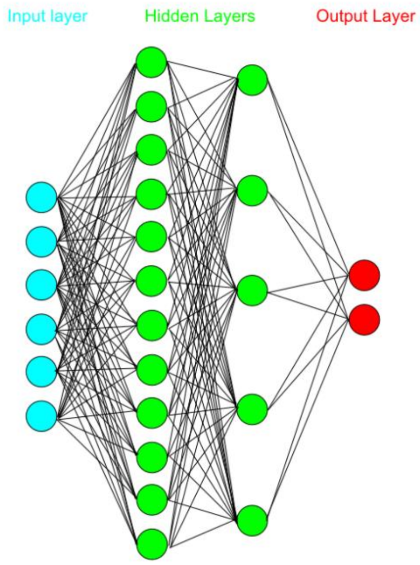

Fig.5. Multilayer Perceptron (MLP)

As the name implies, a Multilayer Perceptron is just that, a network that is comprised of many neurons, divided in layers: input layer, hidden layer or layers and output layer, interconnected by modifiable weights represented by links between layers.

In the input layer number of neurons depends on the number of inputs we want our network to get. In the output layer number of neurons depends on the problem we want the neural net to learn. MLP can have one or more hidden layers. These layers come between the input and the output and their number can vary. The function that the hidden layer serves is to encode the input and map it to the output.

Although MLP can have more hidden layeres, Cybenko has shown that at most one hidden layer is required for approximating functions [12]. However, experience has shown that a two hidden layer network will train significantly faster than a one hidden layer network on some classification problems. But there is still not any specific rule established for selection the quantity of nodes in the hidden layers.

A diversity of training rules have been developed for setting the weights. The most popular rule to date is back propagation which was originally developed by Paul Werbos [13], rediscovered by

David Parker, and popularized by David Rumelhart. The back propagation algorithm is an iterative gradient descent procedure in the weight space which minimizes the total error between the desired and actual outputs of all the nodes in the system.

6 Our face detection system

In this paper we describe the algorithm for locating human faces based on color information and shape analysis based on invariant moments, Hu moments. With the aim to detect human faces we use MLP as a classifier. The method is divided to three steps: skin / non-skin color classification, characterizing clusters using Hu moments and head / non-head classification based on NN.

After processing the original image to binary image, some noise can take place in both skin and non-skin areas. Therefore is necessary to perform opening and closing operations before proceeding to the next step. Opening involves morphological operation of erosion, followed by dilation. By this operation, elimination of noise, made by skin pixel is achieved. To eliminate non-skin pixels noise, closing operation is executed, which involves dilation followed by erosion.

On the resulting image we use the method of invariant moments – Hu moments in order to characterize the shape of each cluster. For that particular image we calculated values of seven Hu moments. Analogously, for each processed image, we repeat the procedure and get the values of the seven Hu moments. Now these values represent the training data set based on which we create NN. All data together are saved as text file in Notepad.

We built our NN in JustNN, neural network system for Microsoft Windows. Creation of neural networks in this system is quite simplified. It allows the user to produce multilayer neural networks from a grid or from text files and images.

We create multilayer neural networks by text file of Notepad. In the input layer of our NN we used seven neurons, seven Hu moments of every image in training data. We only have one hidden layer which has six neurons. One neuron is in output layer. He has a Boolean value i.e. it returns a value true if the NN has recognized a skin patch as head and false value for not-head.

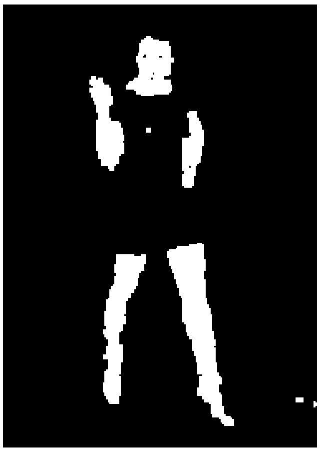

Fig.6. Image after opening and closing operation

7 Experiments

In this section we will present results achieved with our software that utilizes proposed method for face detection. Apart from advantage of simple implementation the method proved to be rather robust. We applied our algorithm on skin patches which present head and non-head and get values of seven Hu moments.

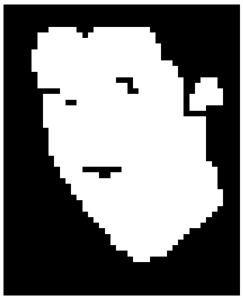

Fig.7. Image of the head

These values were our training data set. Based on them we built MLP. In our net we have 29 training example rows and 24 validating example rows.

From observing and comparing values in tables it is clear that 4th and 5th Hu moments are playing key role in recognizing skin patches as a head.

Based on training set we obtained acceptable recognition success percentages.

8 Conclusion

In this paper, we have presented a new method to detect human faces in color images. We used MLP as classifier of invariant moments after initial detection of skin regions based on skin color in the YCbCr color space. Experimental results show that the proposed method can detect human faces in

ISBN: 978-1-61804-194-4

	h1	*h2*	*h3*	*h4*	*h5*	*h6*	*h7*	*Output*
Picture 1	-0.672599	-200 839	-312 629	-455 734	853 241	560 206	-856 843	1
Picture 2	-0.625782	-171 864	-334 418	-419 423	-804 487	531 498	821 585	1
Picture 3	***-0.63042***	***-168 163***	***-389 065***	***-530 034***	***-105 834***	***666 206***	***990 519***	***1***
Picture 4	-0.208514	-0.456916	-190 752	-195 934	-389 285	-219 171	-563 476	0
Picture 5	-0.106441	-0.246287	-0.736257	-0.79536	-156 119	-0.921291	-351 796	0
Picture 6	0.0258777	0.0055260	-0.864761	-0.783496	-161 222	-0.780936	-244 694	0
Picture 7	-0.060038	-0.193889	-159 001	-287 399	-524 833	-314 226	526 499	0
Picture 8	***-0.639177***	***-16 551***	***-439 554***	***-437 866***	***-876 894***	***-529 276***	***-968 509***	***1***
Picture 9	***-0.665886***	***-190 416***	***-311 183***	***-572 488***	***-104 862***	***-684 694***	***-101 934***	***1***
Picture 10	***-0.582532***	***-169 737***	***-324 228***	***-324 621***	***-665 577***	***-421 963***	***-662 711***	***1***
Picture 11	***-0.507737***	***-124 667***	***-301 022***	***-427 877***	***827 032***	***514 577***	***797 233***	***1***

Table 1. Representation of traning data set with its traning example rows and validating example rows (bold). Hu moments presents inputs while output is 1 or 0 (head or non-head).

color image regardless of size, orientation and viewpoint. Thus the accuracy of this method is quite good, the classification error rate was satisfactory.

References:

[1] M.-H. Yang, D. Kriegman, N. Ahuja. *Detecting Faces in Images: A Survey*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 1, 2002, pp. 34-58.

[2] P. Kakumanu, S. Makrogiannis, N. Bourbakis, *A Survey of Skin-Color Modeling and Detection Methods*. Pattern Recognition, vol. 40, no. 3, March 2007, pp. 1106-1122.

[3] V. Vezhnevets, V. Sazonov, A. Andreeva, *A Survey on Pixel-Based Skin Color Detection Techniques*. GraphiCon, Moscow, Russia, September 2003, pp. 85-92.

[4] J. C. Terrillon, M. N. Shirazi, H. Fukamachi, and S. Akamatsu, *Comparative Performance of Different Skin Chrominance Models and Chrominance Spaces for the Automatic Detection of Human Faces in Color Images*. Proceedings of Fourth IEEE International Conference on Automatic Face and Gesture Recognition, March 2000, pp. 54-61

[5] J.J. de Dios, N. Garcia, *Face Detection based on a New Color Space YCgCr*, ICIP03, 2003.

[6] D. Chai, and K.N. Ngan, *Face segmentation using skin-color map in videophone applications*. IEEE Transactions on Circuits and Systems for Video Technology, vol. 9 no. 4, June 1999, pp. 551-564

[7] R. C. Gonzalez, R. E. Woods. *Digital Image Processing (2nd Edition)*. Prentice Hall, January 2002

[8] G. Kukharev, A. Novosielski, *Visitor identification - elaborating real time face recognition system*. Proceedings of 12th Winter School on Computer Graphics (WSCG), Plzen, Czech Republic, February 2004, pp. 157-164.

[9] M.-K. Hu, *Visual Pattern Recognition by Moments Invariants*, IRE Trans. Information Theory, vol IT-8, pp 179- 187, Feb 1962.

[10] J. Wood, *Invariant Pattern Recognition: A Review*. Pattern Recognition, vol. 29, no. 1, January 1996, pp. 1-17

[11] M. Mercimek, K. Gulez, T. V. Mumcu, *Real Object Recognition using Moment Invariants*. Sadhana Vol. 30, Part 6, December 2005, pp. 765–775

[12] G. Cybenko, *Approximation by superpositions of a sigmoidal function*, Mathematics of Control, Signals, and Systems (MCSS), 2(4):303–314, 1989.

[13] P. J. Werbos. Backpropagation: Past and future. In Proceedings ICNN-88, San Diego, pages 343–353, 1988.