Nipple Detection for Obscene Pictures

PIYABUTE FUANGKHON, THITIPONG TANPRASERT
Department of Computer Science
Assumption University
682 Ramkhamhang Road, Huamark, Bangkapi, Bangkok
THAILAND
piyabute@bis.in.th, nui@s-t.au.ac.th

Abstract: - This paper presents a research on object detection using image processing and neural network, entitled Nipple Detection for Obscene Pictures. Our research work is designed to enhance the effectiveness of pornography detection system by locating nipples in the picture to make sure that the given picture really contains human nudity. Our nipple detection model composing of human skin detection which is used to remove non-skin area and nipple detection which is used to detect nipple on human skin area. Our approach is shown to be effective for a wide range of shades and colors of skin and human configurations. It is validated for locating nipple existence for obscene pictures. In terms of scalability, our approach can be modified to support distributed processing. The system presents excellent performance on a test set of 130 uncontrolled obscene pictures, and a test set of 500 assorted control benign pictures drawn from a wide variety of sources.

Key-Words: - Object Detection, Computer Vision, Erotica/Pornography, Content Based Retrieval.

1 Introduction

A large number of adults and children are using the Internet for searching and browsing though different multimedia documents and databases. To protect the freedom of speech, people are allowed to freely publish various types of material or conduct different types of businesses on the Internet. However, due to the policy, there is millions of obscene pictures and video sequences available for free and commercial download on the World-Wide-Web. Accessing obscene media by children is increasingly a problem with which parents are greatly concerned.

Many research works [12][13][14][15] have been conducted to detect skin and pornography however they address detailed issues regarding the major components of obscene pictures. This points to the question what the major components of obscene pictures are. In our opinion, a picture is considered as obscene if it contains one or more nipples and/or sexual organ. Although the picture may contain very large percentage of human skin area, it should not be concluded an obscene picture unless nipple and/or sexual organs exist in it.

Nipple Detection for Obscene Pictures is an interesting step in obscene picture detection. Given an obscene picture of arbitrary size, our task is to locate the position of nipple in the picture. It is a challenging task since nipples in the picture may appear in different scales and random positions. The imaging conditions, including illumination direction and shadow, also affect the appearance of nipples. Moreover, nipples are non-rigid objects, as there are variations in sizes of nipple base. In addition, the appearance differences among races and genders may considerably add the complexity to the appearance of nipples. A successful nipple detection system should be able to admire the combination of sources of variations.

This paper presents a Nipple Detection for Obscene Pictures. The class/non-class classification problem needs to be addressed because it is not possible to obtain a representative set of non-nipple patterns for training. Furthermore, because of the manifold of sources of variation, a complex decision boundary is anticipated. The classification methods should have a very low false positive rate since the number of non-nipple patterns tested is normally much higher than that of nipple patterns. In addition to a large number of patterns to be tested, a fast classification step is desirable.

Self-Organizing Mapping (SOM) neural network has been widely used in pattern classification. We consider the Nipple Detection for Obscene Pictures using image processing and SOM.

Our algorithm composes of two tasks:

- **Human Skin Detection:** A picture is scanned pixel-by-pixel to determine whether the tested pixel is a human skin by considering the Hue, Saturation, and Value values (HSV color model) of each pixel in the picture.
- **Nipple Detection:** The human skin areas are then fed to the neural network window-by-window to determine whether the tested window is a nipple by using Kohonen's Self-Organizing Maps.

Our model has been validated for detecting nipples in obscene pictures in terms of scalability, which can be modified to support distributed processing. It is shown to be effective for a wide range of shades and colors of skin and human configurations, and presents the excellent performance on a test set of 130 uncontrolled obscene pictures, and 500 assorted control benign pictures drawn from a wide variety of sources.

The structure of this paper is arranged into six sections: Section 2 reviews the background knowledge used in Nipple Detection System. Section 3 proposes the algorithm of Human Skin Detection used to remove non-human skin pixel from the picture. Chapter 4 proposes the algorithm of Nipple Detection used to locate a nipple in the picture. Chapter 5 presents the methodology and experimental results and proposes further tasks. Finally, Chapter 6 summarizes our experiment.

2. Background Knowledge

2.1. RGB to HSV color model conversion

The set of primary colors in the RGB color model has been used in various applications such as image representation and games. However, the RGB color model seems to be not very useful in image understanding. Not only different shades and colors, but also different levels of light or brightness must be considered in the classification of human skin. It is difficult to use RGB color model in classifying human skin. HSV color model [7] [11] is then applied to classify the human skin areas. The color parameters are Hue (H), Saturation (S) and Value (V). Hue is a color attribute that represents a pure color, Saturation defines the relative purity or the amount of white light mixed with a Hue, and Value refers to the brightness of the image.

$$V = \max \tag{1}$$

$$S = \begin{cases} \dfrac{\max - \min}{\max} & if\ \max \neq 0 \\ 0 & if\ \max = 0 \end{cases} \tag{2}$$

$$H = \begin{cases} -1 \times 60 & if\ S = 0 \\ (\dfrac{(G-B)}{(\max-\min)}) \times 60, & if\ R = \max \\ (2 + \dfrac{(B-R)}{(\max-\min)}) \times 60, & if\ G = \max \\ (4 + \dfrac{(R-G)}{(\max-\min)}) \times 60, & if\ B = \max \end{cases} \tag{3}$$

, where max = *sup(R,sup(G,B))* is the maximum value and min = *inf(R,inf(G,B))* is the minimum value of a RGB color model, respectively.

2.2. RGB to Grayscale Conversion

RGB to Grayscale Conversion is applied to quantize the input image from 16.8 million colors to 256 levels of gray colors to enable the estimation of the discrete probabilities. The RGB components of a picture can be converted to Grayscale by using a simple equation as shown in Equation 4 [7].

$$Gray_{(x,y)} = rR_{(x,y)} + gG_{(x,y)} + bB_{(x,y)} \tag{4}$$

$R_{(x,y)}$ is the current value of Red, $G_{(x,y)}$ is the current value of Green, and $B_{(x,y)}$ is the current value of Blue. Based on [08], the value of r=0.299, g=0.587, and b=0.144 is the most appropriate ones.

2.3. Self-Organizing Map

Self-Organizing Map is based on Kohonen Network which is a SOM method having been studied and applied to classify the data by many researchers, such as Teuvo Kohonen [1], Acharya Sushil [2], Ranjit Murhy [3], Janon Ong and Syed Sibte Raza Abidi [4].

The principle goal of SOM is to transform an incoming signal pattern of arbitrary dimensions into one or two-dimensional discrete map, and to perform this transformation adaptively in a topologically ordered fashion [1] [5] [6].

The model of SOM is composed of two layers, input and output layers, each of which is connected via synaptic weights. Input layer can be higher dimensions, but output layer is normally one or two dimensions because the related features of inputs will be demonstrated on the output layer. This paper concentrates on two-dimensional output layer. Figure

1 displays the layout of SOM. The neurons are arranged in a two-dimensional lattice.

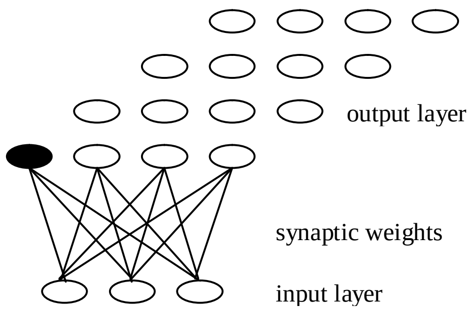

Fig.1 Self-Organizing Feature Map Model

SOM is a competitive learning method. When input vectors are applied to the network, it searches for the winner node. The winner neuron and its neighbors called "activity bubble" will be activated while all others are inactive. To reward the winner and its neighbors, their synaptic weights will be adjusted toward the input vector.

The winner is the cell that best matches with input vector x is found:

$$\|x - w_m\| = \min_i(\|x - w_i\|) \qquad (5)$$

The neighbors of the winning neuron are reduced gradually. The neighbors in two dimensional output layer can be either hexagonal or rectangle.

After the winner neuron m is defined, synaptic weights of a neighbor i is adjusted by:

$$\begin{aligned} &\Delta w_i(t) = a\,(d(i,m),t)[x(t) - w_i(t)] \\ &\textit{for}\ \ i \in N_m(t) \end{aligned} \qquad (6)$$

Where $N_m(t)$ denotes the current spatial neighborhood and $\propto$ is a positive-valued learning function, $0 < \mu(d(i,m),t) < 1$.

Assume that, at the beginning, the output layer has n×n neurons, at the beginning; d is set to be $\lfloor n \div 2 \rfloor$. This means that all neurons are in the neighborhood of the central neuron.

3 Human Skin Detection

Nipple Detection for Obscene Pictures is a system that will locate nipples in a given obscene picture. This picture generally includes both human skin area and non-human skin area such as Tree, Water, and Automobile. In order to increase effectively locate nipple in the picture, we need to remove non-human skin area from the picture thus the detection process will be applied to only human skin area. In this part, we describe our human skin detection algorithm to detect human skin area in obscene picture.

In order to detect human skin pixels, we need to convert the picture from RGB color model into HSV color model because it is easier to define human skin color range using H, S, and V color factors than R, G, and B color factors [7] [8] [11]. After the picture is converted into HSV color model, we try various ranges of H, S, and V color factors that can cover human skin color range. As a result, we found out that the ranges of Hue, Saturation, and Value that can be used to identify human skin are in the ranges [0,36] and [306,360], [0.1,1.0], and [0.1,1.0], respectively. These ranges can be denoted by $H_r \in [0,36] \cup [306,360]$, $S_r \in [0.1,1.0]$, and $V_r \in [0.1,1.0]$. The skin segmentation is defined in Equation 7:

$$Skin_{(x,y)} = \begin{cases} RGB_{(x,y)}\ \textit{if}\ [H_{(x,y)} \in H_r] \cap \\ \qquad\qquad [S_{(x,y)} \in S_r] \cap \\ \qquad\qquad [V_{(x,y)} \in V_r], \\ w \qquad\qquad \textit{otherwise} \end{cases} \qquad (7)$$

The pixels that are classified as human skin are set to the RGB colors. Otherwise, they must be set to white color to exclude them from nipple detection process.

4 Nipple Detection

In our approach, we can generate only a set of nipple windows for the training process because it is impractical to generate a set of non-nipple windows containing everything except nipple. Because of this limitation, we decide not to use supervised neural network like feed forward back propagation that requires us to generate both a set of nipple windows and a set of non-nipple windows. Instead, we decide to implement our nipple detection algorithm based on Self-Organizing Map (SOM), an unsupervised neural network.

In nipple detection, we train our network with a set of nipple windows and use it to detect nipple in human skin area. After the complete training process, a winning neuron for at least one tested data is discovered and considered as valid neuron. The remaining neurons that are not a winning neuron for any tested data become invalid neurons.

In the testing process, the given picture (after we already apply human skin detection) is used to generate a set of testing windows. Each window is composed of a set of pixels. Based on our experiments, 15×15 pixel window is one of the most common size of nipple in the picture. Other nipple sizes are generally larger than 15×15 pixels. Our network tests the set of testing windows and the outputs (winning neurons) are generated. For each window, the output from the network may be the location of valid and invalid neurons. The window is considered to contain a nipple if and only if the output is a location of a valid neuron.

Although the winning neuron may be a valid neuron, the tested window may not contain a nipple at all. This problem can occur when a tested non-nipple window is close to only the winning node. To avoid this problem, the maximum coverage area of each neuron needs to be specified. Any testing window beyond the coverage area of the winning node is not considered as containing nipple.

This stage can be concluded that we add some additional functions to the original Self-Organizing Map to make sure that the winning node represents a valid output. We call this network as Modified Self-Organizing Map (MSOM).

5 Methodology and Experimental Results

Our system operates in two stages: Human Skin Detection (RGB to HSV Conversion, RGB to Grayscale Conversion), and Nipple Detection (Modified Self-Organizing Map). In our experiment, we set the border for all images in order to reduce the size of the testing area. The width and height of the border are 20% of the width and 20% of the height of the image. The border area can be ignored from testing because it was checked earlier that a nipple generally does not exist at the border of the image.

We then separated the picture database consisting of 530 pictures into three sets: a set of 200 obscene pictures (for training), a set of 130 obscene pictures, and a set of 500 benign pictures. The obscene pictures contain one or more nipples of different sizes and positions. The benign pictures contain no nipple. The pictures were collected from various sources including the Internet and in Magazine.

Our network was trained with the following parameters:

1. The network is 20×20 2-dimensional MSOM.
2. The network is in hexagonal.
3. The network is trained 200 epochs.
4. The network is trained with 7,000 15×15 grayscale nipple patterns.

In the testing process, we process the input pictures as follows:

1. Apply Human Skin Detection
2. Apply Nipple Detection to each window into our trained network (Modified Self-Organizing Map)

Our experiments were performed using a PC equipped with 2.8GHz Pentium IV processor with 2GB RAM running MATLAB 7.0. Our testing sets compose of 52,658,235 windows generated from the set of obscene pictures and 189,569,646 windows generated from the set of benign pictures.

To measure the performance of our approach, we use the following measurements:

- A window on non-nipple area detected to contain no nipple is counted as one success.
- A window on non-nipple area detected to contain a nipple is counted as one failure.
- A window on nipple area detected to contain no nipple is ignored.
- One or more windows on the same nipple area detected to contain nipple are counted as one success.

There are two types of input: non-nipple windows and nipple windows. Figure 2 shows the classification of the related data. Figures 3 and 4 show the example of fault detection and successful detection respectively. Tables 1 and 2 present the experimental results for a set of obscene pictures and a set of benign pictures respectively.

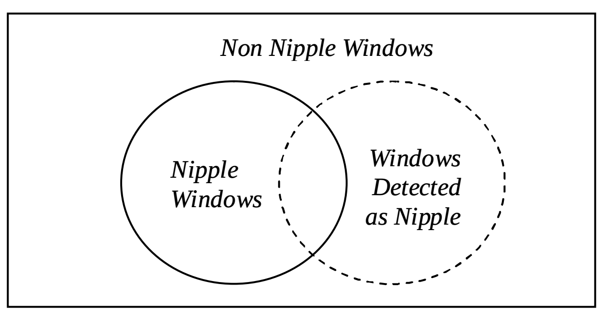

Fig.2 Both nipple windows and non nipple windows are detected as nipple.

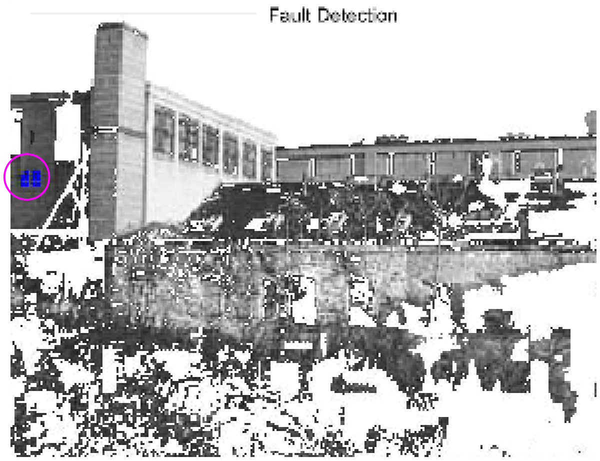

Fig.3 Example of fault detection in benign picture

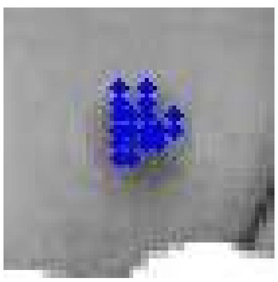

Figure 4: One or more windows containing nipple are detected. This is counted as one success.

Table 1: Experimental result for a set of obscene pictures

	Obscene Pictures
Total Tested Windows	52,658,235
Number of detected Windows	95,899
Total Number of Nipples	445 (100.0 %)
Number of detected Nipples	291 (65.4 %)
Number of none-detected Nipples	154 (34.6 %)

Table 2: Experimental result for a set of benign pictures

	Benign Pictures
Total Tested Windows	189,569,646
Number of detected Windows	409,852
Fault Detection	0.22 %

In our experiment, we collect sample pictures from the Internet and Magazine. We apply human skin detection by using image processing and nipple detection by usingbased on Self-Organizing Maps. Based on the experimental result, if the input is benign picture, the probability of the valid detection is 99.78%. If the input is obscene picture, 65.39% of all nipples in the picture will be detected.

To increase the accuracy of the detection, the dimension of the windows may be adjusted. This will enable the system to detect various dimensions of nipples because the picture may be taken from any camera angel and distance. To increase the detection speed, a user may split the system into human skin detection and nipple detection and distributes them to many computers to speed up the detection process.

6 Conclusion

In order to determine whether the given picture is obscene, it is not enough to consider only the percentage of human skin area in the picture. The approach presented in this paper is able to locate nipples in obscene picture. It can be applied to improve the accuracy of obscene picture detection. We, rather than using supervised neural network, use Kohonen's Self-Organizing Maps, unsupervised neural network, because it is not possible to generate the set of non-nipple pattern for the training process. This paper introduces the uses of unsupervised neural network in classification problem that is normally solved by using supervised neural network.

References

[1] Teuvo Kohonen, *The Self-Organization and Associative Memory*, Berlin: Springer Verlag, 1984.

[2] Sushil Acharya, *Knowledge Discovery and Self-Organizing Systems*, AIT Ph.D. Engineering Thesis, AIT Library, No. CS-97-5, 1997.

[3] Murthy Ranjit, *Knowledge Discovery from Databases using Self Organized Maps*, AIT M. Engineering Thesis, AIT Library, No. CS-95-3, 1995.

[4] Janon Ong and Syed Sibte Raza Abidi, *Data Mining for Ranged Association Rules*, IC-AI'99 International Conference, 1999.

[5] Teuvo Kohonen, The Self-Organizing Map, *Proceeding of the IEEE 78(9)*, pp.1464-1480, 1990.

[6] Simon Haykin, *Neural Networks a Comprehensive Foundation*, Second Edition, New Jersey, Prentice-Hall, 1999.

[7] D. Hearn and M. P. Baker, *Computer Graphics C Version*, Prentice Hall, 1997.

[8] Sanun Srisuk and Weerasak Kurutach, Fast Detection of Scalable and Multiple Human Faces, *InTech'2000 International Conference*, 2000.

[9] Sanhapon Thadapradit and Tang Van To, Using Kohonen network to predict the performance of student at Assumption University based on historical data, *InTech'2000 International Conference*, 2000.

[10] Helge Ritter, Thomas Martinetz and Klaus Schulten, *Neural Computation and Self-Organizing Maps*, Addison-Wesley, 1992

[11] Milan Sonka, Vaclav Hlavac, Roger Boyle, *Image Processing, Analysis, and Machine Vision*, Brooks/Cole Publishing, US, 1999.

[12] James Ze Wang, Gio Wiederhold, Oscar Firschein, *System for Screening Objectionable Pictures Using Daubechies' Wavelets and Color Histograms*, Stanford University, CA, 1997.

[13] D.A. Forsyth, M. Fleck, *Automatic Detection of Human Nudes*, University of California at Berkeley and Claremont, CA, 1999.

[14] David A. Forsyth and Margaret Fleck, *Finding Naked People*, journal reviewing, 1996.

[15] David Brown, Ian Craw, Julian Lewthwaite, *A SOM Based Approach to Skin Detection with Application in Real Time Systems*, University of Aberdeen, UK, 1998.