ISSN 2356-4407

www.STIKI.ac.id

Bersama memajukan ICT Indonesia

PROCEEDING

IC - ITECHS 2014

The 1st International Conference on Information Technology and Security

Malang, November 27, 2014

Published by:

Lembaga Penelitian dan Pengabdian pada Masyarakat

Sekolah Tinggi Informatika dan Komputer Indonesia

PROCEEDING

The 1st International Conference on Information Technology and Security (IC-ITechs)

November 27, 2014

Editors & Reviewers:

Tri Y. Evelina, SE, MM Daniel
Rudiaman, S.T, M.Kom Jozua
F. Palandi, M.Kom

Layout Editor:

Eka Widya Sari

LEMBAGA PENELITIAN & PENGABDIAN KEPADA MASYARAKAT
Sekolah Tinggi Informatika & Komputer Indonesia (STIKI) – Malang
Website: itechs.stiki.ac.id E-mail: itechs@stiki.ac.id

PROCEEDING

The 1st International Conference on Information Technology and Security (IC-ITechs) November 27, 2014

ISSN 2356 - 4407

viii + 276 hlm; 21 X 29,7 cm

Reviewers & Editors:

Tri Y. Evelina, SE, MM
Daniel Rudiaman, S.T, M.Kom
Jozua F. Palandi, M.Kom

Layout Editor:

Eka Widya Sari

Published by:
LEMBAGA PENELITIAN & PENGABDIAN KEPADA MASYARAKAT
Sekolah Tinggi Informatika & Komputer Indonesia (STIKI) – Malang
Jl. Raya Tidar 100 Malang 65146, Tel. +62-341 560823, Fax. +62-341 562525
Website: itechs.stiki.ac.id E-mail: itechs@stiki.ac.id

GREETINGS

Head of Committee IC-Itechs

For all delegation participants and invited guest, welcome to International Conference on Information Technology and Security (IC-Itechs) 2014 in Malang, Indonesia.

This conference is part of the framework of ICT development and security system that became one of the activities in STIKI and STTAR. this forum resulted in some references on the application of ICT. This activity is related to the movement of ICT development for Indonesia.

IC-Itechs aims to be a forum for communication between researchers, activists, system developers, industrial players and all communications ICT Indonesia and abroad.

The forum is expected to continue to be held continuously and periodically, so we hope this conference give real contribution and direct impact for ICT development.

Finally, we would like to say thanks for all participant and event organizer who involved in the held of the IC-Itechs 2014. We hope all participant and keynote speakers got benefit from this conference.

ISSN 2356 - 4407
The 1st International Conference
on Information Technology and Security

LIST OF CONTENT

Object Recognation Based On Genetic Algorithm With Color Segmentation

Evy Poerbaningtyas, Zusana E. Pudyastuti
College of Informatics And Computer Engineering Indonesia
evip@stiki.ac.id, zusana@stiki.ac.id

Abstract

To be able to sort out the optimum color on an observed object is something that is very important, for example in detecting the tumor. An area which has already exposed to the tumor can be identified from the photo. Doctors typically analyze and predict the result of a photo based on their experience, but the doctors can read the photo without predicting the outcome when they use a tool that is able to distinguish color degradation. Thus, creating an artificial intelligence system can help in making a decision. A combination of generic algorithm and image processing is used to optimize color segmentation in detecting an object. In the case of color segmentation of an object, it was found that the spectrum of each color is different one to another. The color differences could be clear, unclear, or almost similar which can be used in detecting and grouping the color differences optimally.

Keywords*: Genetic Algorithm, image processing, color segmentation*

1. INTRODUCTION

The software for solving a real-life problem is increasingly needed, including a software for processing image, since the number of software users are increasing. Image processing applications are well-know for solving problems in medical field in detecting cancer and human blood type, or color recognation of a specific object.

The purpose of this study is to recognize an object in an open area especially in sorting the color in the objects. There are many factors affected the color of the object to be recognized, especially the acceptable lighting. In certain circumstances, the object looks bright when the sun shines brightly and when the weather is cloudy or at dusk, when the light received is reduced. It greatly affects the colors in the object recognition.

The use of genetic algorithm is as a method to determine the best optimization for a number of data. Usually this algorithm is used to solve problems that require an optimum solution. In particular, genetic algorithm can be applied to solve complex optimization problems, particularly in processing color images. This study will discuss about the use of genetic algorithm in optimizing color segmentation. In case of complex optimization, it is expected that there will be a change in the color of an object being observed due to the lightning influence in recognizing an object.

2. THE PROPOSED METHOD

2.1. Digital Image

A digital image can be obtained from an analog image, e.g. an image derived from analog camera/photo). A digitalization process is required to convert an analog image into a digital image. Digitalization process consists of several stages such as sampling and quantization. Sampling process is a process of digitizing a spatial coordinates (x, y) image, while the

quantization process is the process of digitizing the amplitude of analog image to have an intensity function (x, y).

In other words, a digital image visually represented the spatial distribution (2 dimensions) of a physical quantity such as luminance and spatial frequency of an object. Information received might be represented by components such as brightness, color, and so forth. Intensity function (luminance) of spatially distributed objects can be expressed in the form of:

$$f(x, y) = \int_0^{\sim} i(x, y, \lambda).V(\lambda)d\lambda \text{ .....(2.1)}$$

where in:
i(x, y, λ) is the light energy distribution of an object (to express the intensity of a particular light wavelength [λ]) and can be expressed again in the form of:

$$i(x, y, \lambda) = \rho(x, y, \lambda).L(x, y, \lambda) \text{ .....(2.2)}$$

where the ρ (x, y, λ) has a function to state the light reflexivity or transitivity of the object, while the function L (x, y, λ) is the light energy distribution from a light source that illuminates the object.

Light is an electromagnetic wave radiation which can stimulate visual response. The wavelength of light (λ) that appears to the human eyes lies in the visible region, i.e. 350 nm to 780 nm of the electromagnetic spectrum. Luminance range for which the human visual system can work well is approximately 1 to 10 or 10^{10} order of its magnitude. V (λ) is called as luminance relative efficiency function (spectral characteristics) for the visual system.

V (λ) is a relative measurement which does not have a unit. This function arises because of the brightness of 2 kinds of light with 1 (λ 1) and l (λ 2) seen by the human eyes were generally different with $\lambda 1 \neq \lambda 2$ though l (λ1) = l (λ2). To the human eye, an image function of V (λ) was a curve bell-shaped as shown in figure 2.1. below:

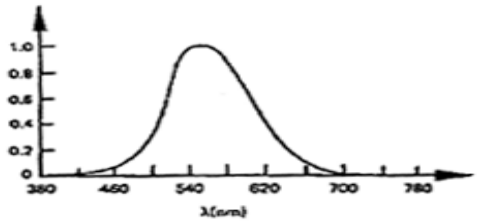


Fig. 1. Brightness level

2.2. Color

An image does not only consist of black and white (bicolor) or the gray level, but also various colors. A colored image will be more interesting than the gray image or black or white image because a colored image will provide more visual information than an image with gray level. Human eyes are able to distinguish thousands of colors rather than to distinguish hundred level of gray.

As we know, color is a reaction of the changing of the light wavelength which can be recognized by the human visual system. Each color has its own wavelength. Of the color spectrum, red could be seen brighter because it has the greatest wavelength, while violet has various wavelength that gives different reactions.

In studying the color, a color usually consists of three basic color components that provide the widest spectrum such as red, green and blue. These three basic colors (often referred to as

RGB) could be formed a variety of colors, but the actual mixing of these three colors cannot cover the whole area of color.

To standardize the system, CIE gives the standard of wavelengths to the three basic colors, as follows:

- Red = 700 nm
- Green = 546.1 nm
- Blue = 435.8 nm

Except the representation of three basic RGB colors, another way to describe the characteristics of an object color image is by representing the color based on its three basic attributes such as *brightness*, *hue*, and *saturation*. *Brightness* represents seen luminance. *Hue* states the degrees of redness, greenness, and so on. In the monochromatic light source, hue difference lies in the different wavelengths. *Saturation* shows the changing perception aspects of white light when it was added to the monochromatic light.

Actually, the above definition is not precise, because the brightness, hue and saturation turn out good wavelength, intensity, hue, and some white light changes.

2.3. Digital Image Processing

There are several image processing operations such as:

1. *Image restoration*
   It is a process for processing digital images obtained closer to its original image. In other words, it is a process to get an original image of a degraded image.
2. *Image enhancement*
   Contrasts to the image restoration, image enhancement is a process of an image to get a more suitable image than the original image for a specific application in which not all types of certain application processing for improving its original image quality is suitable for all existing applications. Thus, specific processing applications are needed to produce suitable image.
3. *Image registration*
   Interpretation of an object is done by a number of images of the object taken separately.
4. *Data image compaction*
   The main objective of image data compaction is to reduce the quantity of data required to represent the data at the storing time on the storage media.
5. *Image segmentation*
   The purpose of image segmentation is to sort out an image into multiple areas or small "important "parts based on a certain criteria. Unlike the previous processing, this process emphasis more on the problem of pattern recognition.

2.4. Genetic Algorithm (GA)

Genetic algorithm is a branch of evolutionary method used to search a value in the optimization problem. This algorithm is taken from the existing genetic processes in living organisms. It refers to the development of generations in a population following the principles of Darwinian natural selection where a strong individual will survive and the weak will extinct. Genetic algorithm imitates the above principles to seek solutions to problems in the real world (Goldberg, 1989).

Genetic algorithm was first proposed by John Holland, which was presented as an abstraction of biological evolution and gave a theoretical mathematical framework for adaptation. Holland's genetic algorithm is a method for moving from one population of

"chromosomes" (bit string that represents an alternate solution to a problem) to a new population of solutions using selection along with a set of genetic operators of "crossover" and mutation. Each "genes" chromosome (i.e. bits) represents an "alleles" particle (i.e. 0 and 1).

General Structure of Genetic Algorithm

Genetic algorithms work on a population which is a set of randomly generated solutions. Each member of the set represents a solution to a problem called chromosomes. The chromosomes in the population evolved in an iteration is called generation. In every generation, each chromosome is evaluated based on an evaluation function (fitness function). The chromosomes are then selected according to their fitness values. Strong chromosomes have a higher probability to exist in the next generation, although there is a probability for weak chromosomes to exist. Of the selection process, a new chromosome (offspring) is formed through the crossover and mutation process from the previous selected chromosomes (parents). Crossover and mutation over the new generation is formed of the selection process.

Genetic Algorithm Cycle above is defined in the following steps:

1. Prepare the initial population.
2. Check the chromosome feasibility when necessary.
3. Calculate the evaluation value (fitness) of each chromosome.
4. Scale the fitness value.
5. Select chromosomes to form a new generation.
6. Conduct crossover and mutation operations.
7. Establish population for the next generation.
8. The selection process and the best chromosomes are used to determine the color being observed when the selection process is convergent or the number of the generations is equal to the limit. If the process does not work properly as expected, the process will start again from step 2.

Characteristics of Genetic Algorithms

In the process of finding the optimum solution, there are two things to consider, i.e. the searching area of exploration and exploitation to find the best solution. Comparisons between hill-climbing algorithm, random searching and genetic algorithm according to the results of the study are as follows (Michaelwicz, 1996):

1. In the hill-climbing algorithm, searching process merely exploits the best solution without giving any attentions to the area of exploration.
2. In the random searching algorithm, the searching process merely explores the searching area without exploiting the best solution.
3. Meanwhile, searching for best solution is carried out optimally by balancing the searching exploration and best solution exploitation in the genetic algorithm.

Natural Genetic	**Genetic Algorithm**
Chromosome	Determined string segment of individual
Gen	Part of string
Loci	Position of gene
Allele	Inserted value of gene
Phenotype	Final solution string
Genotype	A number of strings as the results of matching potential solution

Table 1**.** Comparison between natural genetic with genetic algorithm

In the conventional algorithm, searching the solutions is started from one initial value before the algorithm is exploited through a series of iterations. Meanwhile, the genetic algorithm uses population as its initial value in the searching process instead of one initial value. The iteration of both methods is shown in the image below.

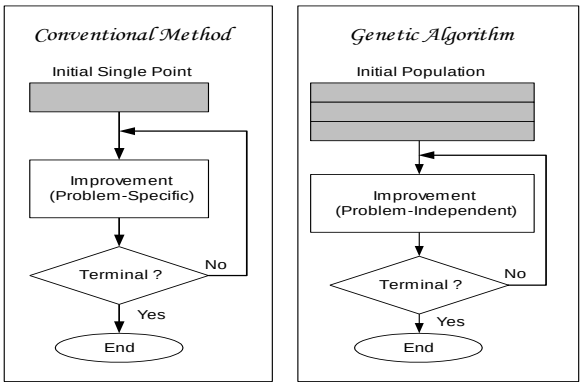


Figure 4. Comparison between Conventional Method and Genetic Algorithm [7]

2.5. Clustering Analysis

In this study, it is necessary to divide the individual members of the sample and the members of the population into several groups in which their characteristics can be expressed into a meaningful designation. Cluster analysis is a data analysis technique which aims to separate / sort the individuals (objects) into groups using their different characteristics, so that the individuals (objects) within a group will have similar characteristics. Clustering is usually performed based on the similarity of the object.

To group *n* individual into *k* groups of $k < n$ is based on the similarity measure of the size indicated by the Euclidean distance. Euclidean distance of X and Y individual in *p* dimension is as follow:

$$D_{(X,Y)} = \left[\sum_{i=1}^{p} (X_i - Y_i)^2 \right]^{1/2} \quad \text{.....(2.5)} \qquad \text{.....(2.16)}$$

D = Euclidean distance of X and Y variables
X_i = X distance
Y_i = Y distance

The greater the D value, the greater the distance between two similar individuals and vice versa. Assumptions that must be met in Euclidean distance are:

- The variables used are not correlated
- Each variable has the same measurement scale

There are two types of grouping in the clustering analysis, namely:

1. Non-hierarchical clustering method
2. Hierarchical clustering method.

Non-hierarchical Clustering Method

In the non-hierarchical clustering method, the amount of *k* has already been known before grouping all the objects / individuals into groups ($k \leq n$). In this case, the distance matrix is not specified and the initial data is not stored during computer processing, so this method is very useful for many cases and is very appropriate in this research. This method starts from the selection of the *k* value as the center of the initial group. By selecting *k* value of the initial group as a free group, one way to choose *k* randomly can be delivered throughout the data or points that are far apart.

One form of non hierarchical method which is often used is *K means*. *K means* clustering method can cluster objects so that the movement of each object to the center cluster within the cluster is minimum.

Step-by-step process of *K means* clustering method is as follows:

1. Divide each item into a number of initial clusters.
2. Process the items listed by determining an item to a centroid (mean) cluster nearby. Then, recalculate the centrid of a cluster which lose an item.
3. Repeat step two until there are no new items enter the cluster.

Hierachical Clustering Methods

Clustering method is a method in which the number of groups that will appear is unknown. It requires several steps, namely:

1. Start clustering from the N cluster, where each cluster contained a single object and calculated the distance between each object, thus *NxM* distance matrix is obtained.
2. Look for the smallest distance values of the two clusters, for example d_{DV} distance is obtained of the U and V.
3. When the $d_{(UV)}$ is the smallest distance, the U and V are combined and given a new label (UV). Then, reassemble the new distance matrix by removing rows and new columns (UV) and rearrange the distance matrix.
4. Repeat step 2 and 3 until all objects are in one group and there are no bias when they are regrouped.

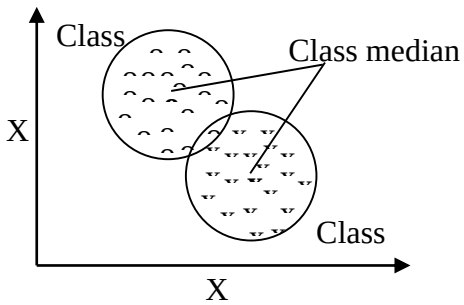


Figure 3. Classification based on the median

3. RESEARCH METHOD

In planning a software and facilitating the organization of a program, a system block diagram and flowchart were made for the overall system to be done in order to ease the

analysis and organizing the programs. Histogram, filter and GA process were included in the system diagram overview. Several sub programs were made in these units to facilitate the organization of overall software better

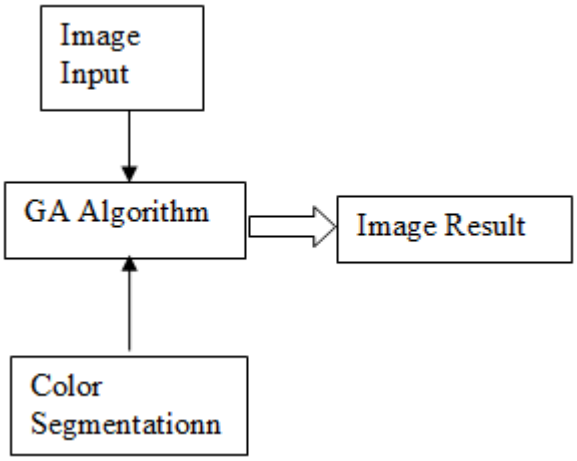


Figure 4. System Diagram

The stages of system design were as follow:

1. The picture was taken using an image file. Once the image was displayed, the colors to be observed were determined.
2. After the image taking and color determining, the next stape was to input both of the parameters into GA process. This process had several procedures.
   a. Determining the number of initial population,
   b. The population generated randomly and determined the individual,
   c. The determination referred to the image to be processed. Each individual consisted of several parameters, namely the R value, G value, and B value. The individual determined as fitness value when the total value of RGB met the parameters or closed to the expected criteria. To find the fitness value, the selected individuals experienced the crossover process before obtaining the new individuals. It also experienced the mutation process before determining the fitness value based on the image. These process was an ongoing process through several generations.
   d. To ensure the fitness value at each iteration to remain better, the elitisms process was conducted. Thus, the value resulted at each process would always be better than previous or at least it would have an equal value to the previous iteration value.

Figure 5 shows the GA Process where the colors were groups based on their fitness value determined during the GA process.

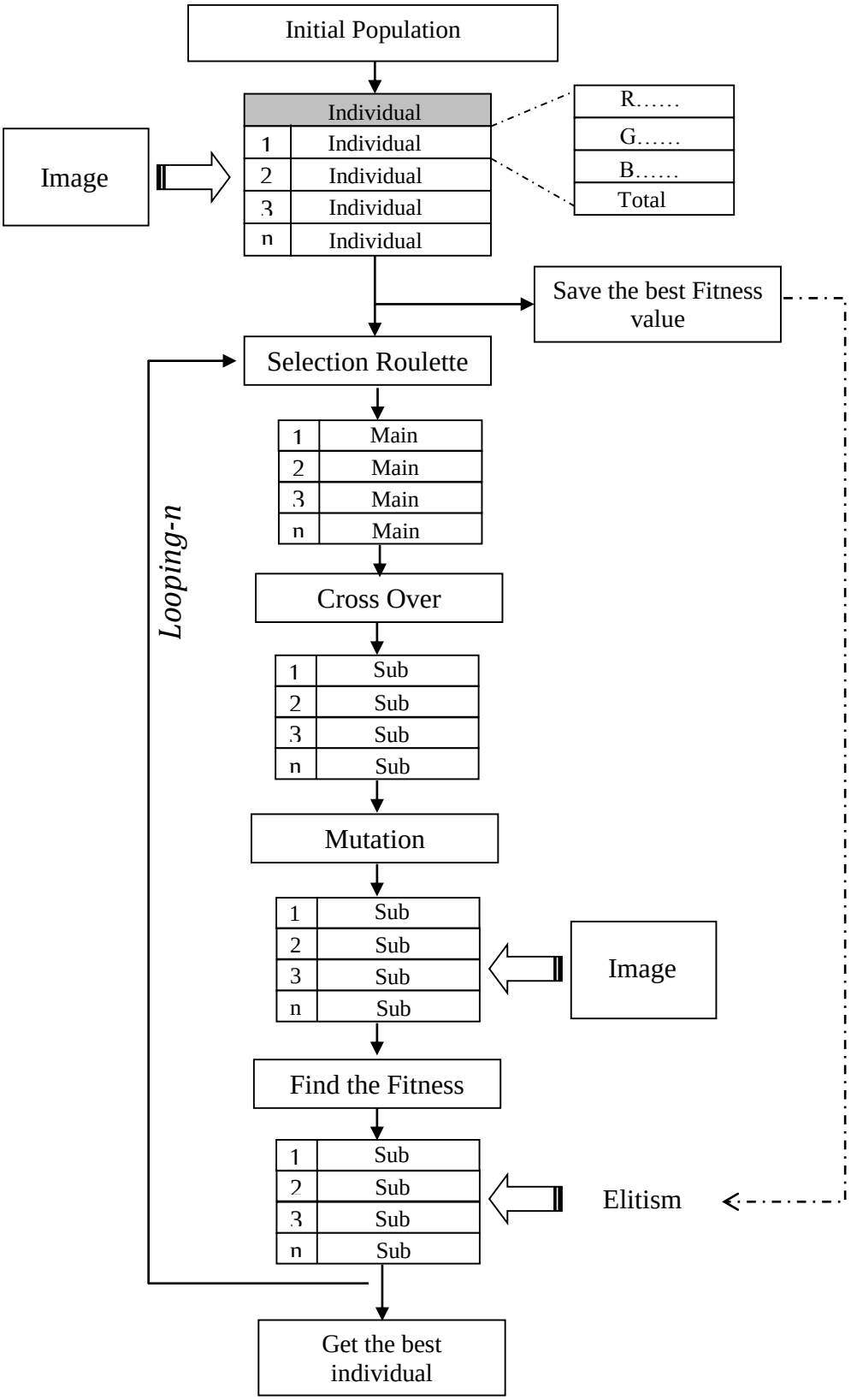


Figure 5. Diagram Alg. GA

4. RESULT AND DISCUSSION

The software was tested many times by providing a wide range of inputs to have an expected software or to minimize the errors. The software was also tested without the use of cameras (off line), and the data to be processed came from the existing image file.

The object to be tested was an image of a baby who was eating some food. Object recognition using color sorting was used to find out what object was being eaten by the baby.

Figure 6. Tested object

The testing stages taken were:

First, the population value was generated to find the greatest value of the population of the object (Figure 6). For instance, when the number of generations was 25, the greatest value of these 25 generation was selected for the next process.

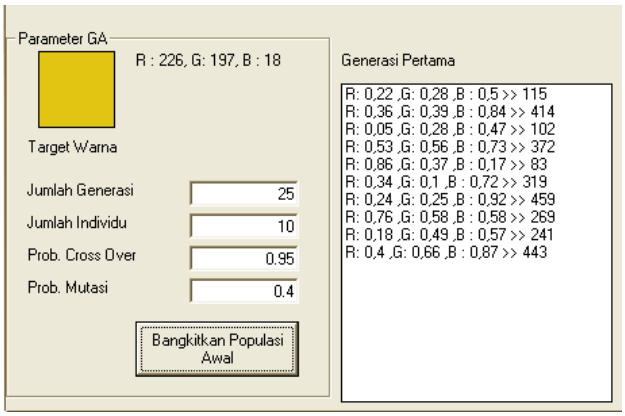


Figure 7. Process of generating population values towards the R, G, B colors

Second, color segmentation was performed by clustering the color and applying the GA algorithm to find the best fitness value for color segmentation of the searched target. The threshold value referred in this stange determied the GA.

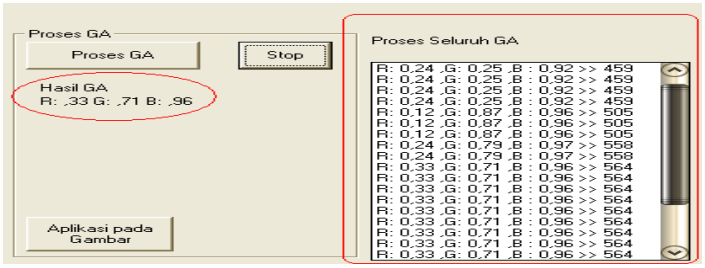


Figure 8. Color segmentation Process

Next, input the threshold value to group the color of the search target in which the result were displayed in the following images:

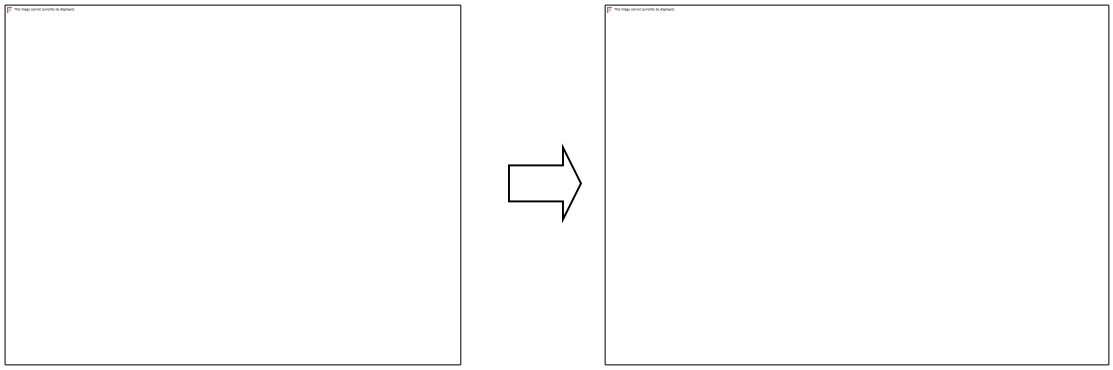
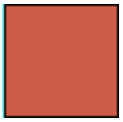
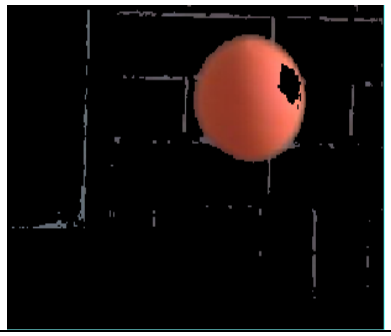


Figure 9. Image of testing process

Of the previous clustering process could be seen the ability of the system (GA) to group the color by replacing the value of the generations number or individuals that certainly affected the outcome of a color grouping.

The clustering process above can show how far the clustering process can change the values of a number of generations or individuals which affected the result of color segmentation. The results of color segmentation to some objects are as follow.

OBJECT IMAGE	TARGETED COLOR	RESULT IMAGE
		
		

Tabel 2. The test result of several objects

The results of offline color segmentation (using image capture file) gave good results because the image taken for a sample was an image file with very little noise. Thus, the capturing distance of an object, the level of intensity around an object and the image quality

used in object greatly determined the outcome of the colors segmentation. The ability of a software to segment the color was recognized by knowing the results.

5. CONCLUSION

Based on the analysis of the experiments, it could be concluded that the system has already provided good results to develop an online system:

1. In the experiment, when there was a little spectrum differences between the targetted color to be segmented and there were a little amount of background color, the result expected contained errors. It is shown on Table 2. The success rate was about of 40%.
2. In searching for the fitness value, the color segmentation had better results when the number of generation and populations was greater, but the process would be slower. It made the determination of generations value and the population value as a very important issue because the results obtained and a short processing time must be taken into a consideration.
3. The distance of objects captured, the level of intensity around the objects and the quality of the camera used when capturing the objects determmined the outcome of the colors segmentation. Thus, the capabilities of the software in this case could be recognized as shown in Table 2.

DAFTAR PUSTAKA

[1] Achmad Basuki, *"Algoritma Genetika (Suatu Alternatif Penyelesaian Permasalah Searching)",* PENS-ITS. Surabaya 2003.

[2] Guarav Sharma *"Digital Color Imaging Handbook"* Xerox Corporation Webster, NY. CRC Perss 2003.

[3] Hu, Xiaochun. 2003. "*Comparison of Genetic Algorithm and Neural Network Technology"*. http://www.ics.hawaii.edu/~xiaochun

[4] I Made Wiryana, SSi, SKom, MSc "*Penggunaan Metoda Soft Computing Untuk Aplikasi Bisnis*". Universitas Gunadarma. Jakarta 2003.

[5] Jani F Mandala. "*Pemanfaatan Transformasi Wavelet Citra Wajah Sebagai Sistem Keamanan Kunci Kombinasi*", Fakultas Pasca Sarjana – Jurusan Teknik Elektro – Teknik Sistem Komputer – ITB. Bandung 2003.

[6] Jim Bruce *"Color Segmentation"*, 15-384 Computer Vision – Carnegie Mellon University 2003.

[7] Mainzer, Tomas. "*Genetic Algorithm for Traffic Sign Detection"*. University of West Bohemia. Czech Republic.

[8] Michalewicz, Zbigniew.1996. *"Genetic Algorithms + Data Structures = Evolution Programs"*, 3rd rev. and extended ed, New York: Springer – Verlag Berlin Heidelberg.

[9] Mulyanto, Eko. Arief Kurniawan, Warsono. "*Perencanaan dan Pembuatan Perangkat Lunak Perbaikan Citra Digital Menggunakan Algoritma Genetik".* Proceeding of Seminar on Intelligent Technology and Its Applications (SITIA 2002) Jurusan Teknik Elektro FTI-ITS Surabaya

[10] Professor G.D. Hager "*Computer Vision Segmentation*". http://www.cs.jhu.edu/~hager

[11] Real-time color segmentation of road signs (*Mohamed BénallalGraduate "School of École des Mines de Paris", Jean Meunier"Department of Computer science University of Montréal")*

[12] Sri Kusumadewi, *"Artificial Intelligence (Teknik dan Aplikasi)*".Jogyakarta Graha Ilmu. Jogyakarta 2003