31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

A New Algorithm for Segmenting and Counting *Aedes aegypti* Eggs in Ovitraps

G. Gusmão, Saulo C. S. Machado, and Marco A. B. Rodrigues

Abstract—**Dengue fever has become a major international public health concern in recent decades. As dengue fever not have available vaccine or specific treatment, the only known form to prevent the illness is by applying strategies to control its vector, the *Aedes aegypti* mosquito. Ovitraps, special traps to collect mosquito eggs, are used to detect *Aedes aegypti* presence and to approximate the gauge of the adult mosquitoes population in the environment by counting the number of eggs laid in an trap. This counting is usually performed in a manual, visual and non-automatic form. This work proposes a new automatic method to automatically count the number of eggs in digital images of ovitraps based on image processing techniques (color systems exploration) and k-Means clustering algorithm. The proposed method performs an improvement on the results when compared with previous studies.**

I. Introduction

DENGUE fever is currently the most globally widespread insect-born virus infection, causing 50 - 100 million cases per year in more than 100 endemic countries. Dengue fever is found in tropical and sub-tropical regions around the world, predominantly in urban and semi-urban areas. Dengue virus is transmitted to humans mainly by *Aedes aegypti* mosquito, which is also vector of yellow fever.

The first historic case of dengue fever in the world was recognized in Java Island in 1779. In Americas, the disease was related over 200 years ago, with epidemics in Caribbean and United States. In 1982, the first Brazilian epidemic was confirmed by laboratorial tests in Boa Vista (State of Paraná), although it had had indications of epidemics since 1923 [1].

The spread of dengue fever is attributed to the expanding geographic distribution of the four dengue viruses and their vectors. The rapid rising of the urban mosquito population is bringing even higher number of people into contact with this vector, especially in areas that are appropriate for mosquito breeding, e.g., locals were where household water storage is common and where solid waste disposal services are inadequate [2].

N. M. Portela, C. A. B. Mello, W. P. dos Santos, C. M. G. Gusmão and S. C. S. Machado are with Polytechnic School of Pernambuco, University of Pernambuco, Recife, Brazil, 50720-001 (e-mails: {nmp, wellington, cristine}@dsc.upe.br, carlos@ieee.org, Phone: +55 81 3184-7542).

M. A. B. Rodrigues is with Department of Electronic and Systems, Federal University of Pernambuco, Recife, Brazil, 50670-000 (e-mail: benedetti@ufpe.br).

A. L. B. Candeias is with Department of Cartographic Engineering, Federal University of Pernambuco, Recife, Brazil, 50740-530 (e-mail: analucia@ufpe.br).

As dengue fever does not have available vaccine or specific treatment yet, the only form to prevent the illness is to apply strategies of vector control which demand that areas of risk and periods of risk are identified [3]. Entomological surveillance is used to indicate priority areas to allow stratification of control measures and to help the decision of which control strategies will be used at a certain time. The primary purpose, however, should be to identify areas and periods of time when it is most probable dengue fever occurrence [4].

Dengue vector surveillance is classically based on the Premise Index and the Breteau Index, both of which use visual detection of larvae in domestic containers. *Aedes aegypti* larvae visualization is an inaccurate technique because of the larvae's ability to escape rapidly and their capacity to remain submerged for long periods of time. The percentage of premises or containers where *Aedes aegypti* larvae are found does not provide information regarding the mosquito population density (registering as positive a container whether just one or thousands larvae are present,). These indices do not seem to be an adequate way of meeting vector surveillance needs [5].

Ovitrap surveys could be considered an effective and efficient technique for detecting and monitoring *Aedes aegypti* populations at low densities [6]. Using ovitraps for vector surveillance seems to be a current trend in dengue endemic countries, since this method is more sensitive and it allows better assessment of infestation densities than the conventionally used methods based on the search for larvae [5, 7, 8].

Next Section describes the scenario which this study was performed, the images acquired and the new segmentation algorithm developed for the automatic counting of *Aedes aegypti* eggs in ovitraps. After this, the results are presented and analyzed in Section III. Section IV concludes the paper and relates some commentaries on our results.

II. A Geographic Informational System for Dengue Control

Health Technology Assessment should inform health decision-makers about the introduction of new healthcare technologies. Challenging image acquisition and analysis problems require a unique mix of hardware support, analysis capabilities, and image processing algorithms. In these situations, scientists and engineers need an open and flexible architecture to create a customized workflow without wasting time on low-level details.

978-1-4244-3296-7/09/$25.00 ©2009 IEEE

A Geographic Information System (GIS) is an automated computer-based system with the ability to capture, retrieve, manage, display and evaluate large quantities of spatial and temporal data in a geographical context. The system comprises hardware (computer and printer), software (GIS software), digitized base maps, information and a whole set of procedures such as data collection, management and updating.

Specific diseases and public health resources can be mapped in relation to their surrounding environment and existing health and social infrastructures. Such information when mapped together creates a powerful tool for the monitoring and management of disease, in this case dengue fever. GIS provides a graphical analysis of epidemiological indicators over time - provided by counting the number of eggs in ovitraps, captures the spatial distribution and severity of the disease, identifies trends and patterns and indicates where there is a need to target extra resources.

Ovitraps consist of black containers that are partially filled with tap water holding vertical wooden paddles with one rough side [9]. They are safe, economical and environment-friendly. In regions of limited manpower for vector control, the extensive use of ovitraps is an important resource to help collect data on *Aedes aegypti* population on a wider area and gauge the effectiveness of control efforts [10].

Currently, ovitraps are used as a means of detecting *Aedes aegypti* presence as well as an approximate gauge of the adult population in an area. It can be used to estimate fairly well the population of adult mosquitoes in the environment by counting the number of eggs laid on the moist paddle [10].

This counting is usually performed in a manual, visual and non-automatic form. To aid the control of dengue proliferation and the eradication of its more dangerous form to human beings, dengue haemorrhagic fever, this work approaches the development of automatic methods to count the number of eggs in ovitraps images using image processing and analysis, particularly color systems exploration k-Means clustering technique.

An ovitrap layer comprising a spatial map and an attribute table was created in the GIS for monitoring and evaluating the network of ovitraps placed citywide to better understand vector trends and disease patterns. To every ovitrap placed is given a unique identification number and its spatial location is stored in the GIS. A table is written behind the ovitrap which preserves the ovitrap's criteria of identification (e.g., number, the date of the weekly collection, the address of the site, the position of the ovitrap, the species found in the ovitrap, the larval instars and pupal stages, etc.) and the dominant species type breeding in that ovitrap for that week.

In this scenario, this study aims to bridge a project for developing a GIS to evaluate the ovitrap breeding data collected weekly to identify hotspots and risk areas where there is a danger of high *Aedes aegypti* infestation. The analysis results will be used to support health decision-makers in planning vector surveillance and control operations.

In Recife, Brazil, most of the research projects involving dengue infestation mapping by using Geographic Information System, surveillance strategies and dengue proliferation control are performed in conjunction with Aggeu Magalhães Research Center (CPqAM). This research is part of a project called SAPIO, granted by FINEP, that aims the development of new technologies for dengue control, surveillance and information dissemination [11].

III. Materials and Methods

A. Acquisition of images

For digitization of the ovitraps, a digital camera was used with the following parameters: 7.2 Megapixels resolution, LCD 2.5'', 4.5 times Optical Zoom and LEICA DC Vario Elmarit lens. The ovitraps were digitized with about 700 dpi resolution and 4 times optical zoom. This process generated true color digital images of 3,072 versus 2,304 pixels. The same image was split into 6 sub-images for the experiments. The amount of eggs in each one of these sub-images is acquired by visual inspection performed by technical specialists.

B. K-means clustering

Initially, the images are acquired in RGB color system and then they are converted into L*a*b* color system [12]. The component L* indicates luminosity while the components a* and b* indicate chromaticity information. Fig. 1 (top-left) presents a RGB sample image and the components of its conversion to L*a*b*: L* (top-right), a* (down-left) and b* (down-right). The L*a*b* color system derives from the XYZ color system. Its conversion is given as follows [12]:

$$\begin{aligned} L^* &= \begin{cases} 116(Y/Y_n)^{1/3} - 16, & \text{if} \quad (Y/Y_n) > 0.008856 \\ 903.3(Y/Y_n), & \text{if} \quad (Y/Y_n) \le 0.008856 \end{cases} \\ a^* &= 500\left[f(X/X_n) - f(Y/Y_n)\right] \\ b^* &= 200\left[f(Y/Y_n) - f(Z/Z_n)\right] \end{aligned} \tag{1}$$

where X_n, Y_n and Z_n are white reference and $f(t)=t^{1/3}$, if $t > 0.008856$, otherwise $f(t) = 7.787t + 16/116$.

Components a* and b* are used to cluster the input image using k-means map [13]. For this experiment, the number of clusters was set to three. Such clusters are associated to the following classes: egg, trap and intermediate regions. The proposed algorithm randomly selects the initial mean vectors. The main target here is to minimize the sum of Euclidean distances from each object to the cluster mean vector (centroid) they are associated to, for each cluster in the total set of clusters. This measure is called cohesion. As stop criteria, there were used two measurements: a maximum of iterations of 100; and the cohesion estimative. The algorithm stops when the sum of distances for each object cannot be decreased further. The algorithm repeats the clustering 3 times, each with a new set of initial cluster mean vectors positions, and returns the solution with the lowest value for the sums of point-to-mean vectors distances. The results of k-Means clustering can be seen in Fig. 2.

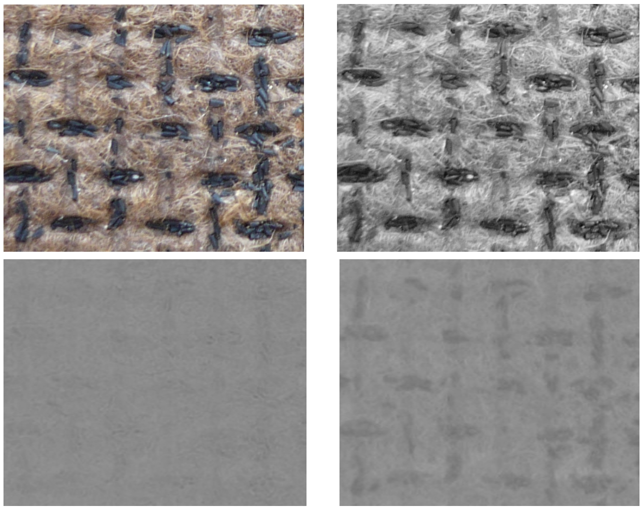

Fig. 1. (top-left) RGB color image of an Ovitrap and the L*a*b* components of the image after its conversion: (top-right) 'L*' luminosity, (bottom-left) chromaticity component 'a*' and (bottom-right) chromaticity component 'b*'.

Fig. 2. An ovitrap image clustered using k-Means clustering technique.

C. Image labeling

After clustering the images, it is necessary to define which cluster contains eggs. For this purpose, the original RGB image is converted into a HSV (H = hue, S = saturation, V = value) image. Hue is related with the tones that the image contains. Thus, an analysis of the hue component can give us information about the objects. The hue is evaluated as follows [12]:

$$H = \begin{cases} 60(G-B)/(M-m), & if \quad M = R \\ 60(B-R)/(M-m), & if \quad M = G \\ 60(R-G)/(M-m), & if \quad M = B \end{cases} \tag{2}$$

where R, G and B are the values of the red, green and blue components for a given color, m is the minimum value between R, G and B and M is the maximum value between them. For each cluster, it was evaluated the average hue value. It was found experimentally, based on the histograms of each class, that when this value is higher than 0.5, the respective cluster is set an area with eggs (then it turns to 1), otherwise it is ignored (and it is turned to 0). A bi-level image is the result of this analysis as Fig. 3-left illustrates.

A connected components algorithm is applied to the bi-level image in order to label its connected regions [14]. This algorithm puts a different label at each connected white area of the image. With this labeling, it is possible to measure each connected area. Small areas are deleted as it could not contain an egg. By experimentations, it was defined that every area with less than 140 white connected pixels should be erased. This can be seen in Fig. 3-right where it is presented the image of Fig. 3-left after the removal of its white areas. This is done to decrease the noise in the image.

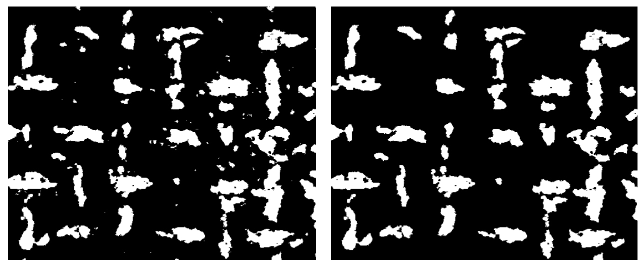

Fig. 3. (left) A bi-level image and (right) the same image after elimination of small connected areas.

Finally, it was considered that an egg occupies an area of 357 pixels. Thus, the number of eggs is the total amount of white pixels divided by this average area. In this case, the method registered an amount of 118 eggs against the correct value of 111 eggs that the image contains.

The proposed method robustness can be seen in Fig. 4 where small rocks present in the images are completely distinguished from the eggs in clustering results avoiding false-positives. It is an important result because by other features (as form and size) a rock could be easily misclassified as an egg by the algorithm.

Fig. 4. (left) Examples of images with rocks and (right) a segmentation result.

IV. Discussion and Results

In Table I, it is presented the results of the method applied to six samples, including an image with no eggs. The image labeled as '3' in Table I is the image previously presented in Fig. 1-top-right with 111 eggs.

The method described herein reached a maximum individual error of 45.45% in the first image where there is a difference of ten eggs. But in general, the average error was about 1.16% and the standard deviation was about 4.15 which is acceptable in comparison with a non-automatic method. The average error is more important than the individual errors as the method will be applied in the complete image in practical use.

TABLE I
RESULTS USING THE PROPOSED METHOD

Image	Correct Amount of Eggs	Estimated Amount of Eggs by the Proposed Algorithm
1	22	32
2	8	9
3	111	118
4	30	20
5	19	20
6	0	0
total	255	258

The proposed method presents an improvement on the average error from 6.66% to 1.16% when compared with previous results achieved by Mello *et al*. [11]. On the other hand, the maximum individual error increased from 25% to 45.45% and the standard deviation increased from 1.6 to 4.15.

V. Conclusion

The *Aedes aegypti* mosquito is the main vector of dengue illness, a major international public health concern in recent decades. The mosquito lay his eggs in ovitraps and the informations about the egg gauge could genarate important statistcs to fight the illness applying strategies of vector control.

Looking for a precise and robust method to replace the current couting process performed in a manual, visual and non-automatic way, the major objective of this work is to propose an algorithm for segmenting and counting *Aedes aegypti* eggs in ovitraps images automatically.

The novel method uses L*a*b* and HSV color systems exploration and k-Means clustering technique. An important feature of the proposed method is that it had a better global result than Mello *et al.* decreasing the average error.

Future works will concentrate on the application of evolutionary computing to perform a clustering parameters optimization by analyzing clustering quality indexes as the quantization error, cohesion between points in the same cluster and distances between clusters.

Acknowledgment

This research is partially sponsored by FINEP-Brazil. The authors are grateful to Aggeu Magalhães Research Center for the concession of the ovitraps used in this study.

References

[1] *Dengue: Instruções para pessoal de Combate Ao Vetor – Manual de Normas Técnicas*, 3rd ed., FUNASA, Brasília, 2001 (2009, April 3) [Online]. Available: http://bvsms.saude.gov.br/bvs/publicacoes/funasa/man_dengue.pdf .

[2] World Health Organization (2009, April 3), *Dengue and dengue haemorrhagic fever*, [Online]. Avaliable: http://www.who.int/mediacentre/factsheets/fs117/en/ .

[3] A. Tran *et al.*, "Dengue spatial and temporal patterns, French Guiana, 2001," *Emerg. Infect. Disease*, vol. 10, n. 4, pp. 615-21, April 2004.

[4] M. R. Dibo *et al.*, "Study of the relationship between *Aedes (Stegomyia) aegypti* egg and adult densities, dengue fever and climate in Mirassol, state of São Paulo, Brazil," *Mem. Inst. Oswaldo Cruz*, vol. 03, n. 6, pp. 554-560, September 2008.

[5] L. Regis *et al.*, "Developing new approaches for detecting and preventing *Aedes aegypti* population outbreaks: basis for surveillance, alert and control system," *Mem. Inst. Oswaldo Cruz*, vol. 103, n. 1, pp. 50-59, February 2008.

[6] K. L. Chan, S. K. Ng, K. K. Tan, "An autocidal ovitrap for the control and possible eradication of *Aedes aegypti*," *Southeast Asian J. Trop. Med. Pub. Hlth.*, vol. 8, n. 1, pp. 56-62, March 1977.

[7] V. C. G. Morato *et al.*, "Infestation of *Aedes aegypti* estimated by oviposition traps in Brazil," *Rev. Saúde Pública*, vol. 39, n. 4, pp.553-8, August 2005.

[8] M. R. Dibo *et al.*, "Identification of the best ovitrap installation sites for gravid *Aedes (Stegomyia) aegypti* in residences in Mirassol, state of São Paulo, Brazil," *Mem. Inst. Oswaldo Cruz*, vol. 100, n. 4, pp. 339-343, July 2005.

[9] R. W. Fay, A. S. Perry, "Laboratory studies of ovipositional preferences of *Aedes aegypti*," *Mosquito News*, vol. 25, pp. 270-281, 1965.

[10] G. T. Ai-leen and R. J. Song, "The use of GIS in Ovitrap Monitoring for Dengue Control in Singapore," *Dengue Bulletin*, vol. 24, pp. 110-116, December 2000.

[11] C. A. B. Mello *et al*., "Image Segmentation os Ovitramps for Automatic Counting of *Aedes Aegypti* Eggs", in *30th Annu. International Conf. of the IEEE Engineering in Medicine and Biology Society*, Vancouver, 2008, pp. 3103-3106.

[12] L.Velho, A.C.Frey, J.Gomes, *Image Processing for Computer Graphics and Vision*. New York: Springer, 2009.

[13] R. O. Duda, P. E. Hart, D. G. Stork, *Pattern Classification*. New York: Wiley, 2001.

[14] L.Shapiro and G.C.Stockman, *Computer Vision*. New Jersey: Prentice Hall, 2001.