

Reprinted from

**Symposium on
Machine Processing of
Remotely Sensed Data**

June 29 - July 1, 1976

The Laboratory for Applications of
Remote Sensing

Purdue University
West Lafayette
Indiana

IEEE Catalog No.
76CH1103-1 MPRSD

Copyright © 1976 IEEE
The Institute of Electrical and Electronics Engineers, Inc.

Copyright © 2004 IEEE. This material is provided with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the products or services of the Purdue Research Foundation/University. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org.

By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

AUTOMATIC DETECTION AND CLASSIFICATION
OF INFESTATIONS OF CROP INSECT PESTS AND DISEASES
FROM INFRARED AERIAL COLOR PHOTOGRAPHS

M. Ali* and J.K. Aggarwal

Department of Electrical Engineering, The University of Texas at Austin, Austin, Texas.

I. ABSTRACT

This paper describes a system designed to analyze aerial color infrared photographs of citrus orchards. The input to the system is a 35 mm transparency of the infrared aerial photograph mounted in a 3-color film digitizer which is controlled by an XDS 930 computer.

An algorithm locates the outline of each individual tree in the photograph and then cluster analysis is applied. For the trees of a transparency, centre of gravity (C.G.) of the set of these trees is determined and then the distance of each individual tree from C.G. is computed by using an empirical distance function which we have derived. Sorting, performed on the basis of the computed distances, clusters the trees. The cluster identifying algorithm separates out the individual clusters which are then grouped together into a smaller number of classes by applying the same algorithm. Fifteen transparencies having mealybugs, brown soft scale, gummosis and root rot infestations, were analyzed and experimental results show very satisfactory classification of the infestations.

II. INTRODUCTION

Detection and control of crop diseases and plant insect-pest infestations are important phases of agricultural management throughout the world. Plant diseases and insect pests have devastating effects on crop yield and reduce the quality and the market value of the products. Damage may begin at the time the crop is planted, continue throughout its growing period, and persist after harvest, when products are transported and placed in storage. Unless pests are prevented or controlled, there can be no sustained improvement in agricultural productivity.

The economic benefits of early detection of disease or infestation of crops are the same as increasing agricultural productivity. Even in developed countries, losses in crop value caused by insects and diseases are high. In the United States, the average annual losses caused by plant diseases are estimated at about \$3.7 billion [1]. Average annual losses from crop insect pests amount to \$3.8 billion.

Peach tree decline problem has been investigated in Georgia since 1929. An estimated 200,000 to 300,000 trees died during the spring and summer of 1962 [2]. During the spring and summer of 1972, a substantial rise in tree decline of the order of 200,000 trees was again observed, which is approximately 10% of the productive capacity of the state. The premature death of peach trees in the southeast has been attributed to such factors as Cytospora Canker, Clitocybe, root rot, peach tree borers, nematodes, bacterial canker, pythiaceae fungi, virus diseases, poor cultural practices and repeated planting of peaches on the same sites [3-16].

The citrus blackfly, *Aleurocanthus Woglumi* Ashby, is a damaging pest of citrus crops and has been a constant threat to the citrus industry in the United States. It invaded the Florida Keys in 1934 [17], south Texas in 1955 [18] and again in 1971. Three years of efforts were required for eradication of the infestation in Key West and one year for the first eradication in Texas.

The phony disease and European red mite also attack peach trees. Phony disease of peach is a virus disease transmitted by five leaf hopper species [19-22].

Five pecan anomalies have been detected. Prionus larvae attain a length of three inches and a weight of fifteen grams. They travel through the soil feeding on and in pecan roots; roots are even girdled and severed by the larvae. Severe feeding by the beetle larvae is associated with gradual limb-by-limb death of pecan trees [23]. Foliage on dying or declining trees appears more irregular, thinner, and a lighter green than foliage on healthy trees.

The characteristic symptom of bunch disease is a bushy growth of slender, willow shoots that result from an abnormally large number of lateral buds being forced into growth [24]. Bunch disease appears to spread from tree to tree by undetermined means. In some respects, symptoms resemble those of severe virus diseases of other woody plants which are spread by mites or insects. The color of the diseased leaves varies from dark green to

*On leave from Computer Centre, Aligarh Muslim University, Aligarh, U.P., India.

chlorotic; the terminal leaflets are subnormal in length and the diseased leaflets are usually thinner and wider than healthy leaflets. The leaflets on diseased shoots are frequently wavy or twisted and appear slightly wilted. Unusually, bunch disease is more easily recognized in the early than in the more advanced stages.

Among forest diseases, *Poria Weirii* root rot is especially destructive. It is responsible for extensive killing of 25 to 125 year old Douglas-fir trees in the Pacific Northwest region of the United States. The disease is estimated to have an annual growth impact that exceeds 100 million cubic feet in the United States.

Langley and Colwell's [25-27] research enhanced the use of aerial photo investigations and particularly stimulated research with infrared films. However, up to 1965 very little use was made of multi-band sensing to overcome the problems of survey and detection of crop disease and insect infestation. Since that time black-and-white, infrared and Ektachrome Infrared photographs have been used experimentally to detect plant diseases. Infrared aerial color photography offers many advantages over ground survey. On color infrared film the information concerning insect infestations, diseases, nutrient deficiencies, soil problems, freeze damage, density and distribution patterns of vegetation and differences in varieties produces color variations, textural differences, patterns of distribution or geometric distortions of the vegetation. Some experiments [28-30] indicate that certain plant diseases can be detected by a remote sensor before they can be detected through visible means.

Photographically recorded foliage reflectance in the near-infrared has been much used because spongy mesophyll tissue of a healthy leaf, which is turgid, distended by water, and full of air spaces, is a very efficient reflector of any radiant energy and, therefore, at the near-infrared wavelengths. These pass the intervening palisade parenchyma tissue (which absorbs blue and red and reflects green from the visible). When water relations are disturbed and the plant starts to lose vigor, the mesophyll collapses, and as a result there may be great loss in the reflectance of near-infrared energy from the leaves almost immediately after the damaging agent has struck a plant. Furthermore, this change may occur long before there is any detectable change in reflectance from the visible part of the spectrum, since no change has yet occurred in the quantity or quality of chlorophyll in the palisade parenchyma cells. Therefore, to detect this change photographically, a film sensitive to the near-infrared wavelengths is used. Since infrared-sensitive films are also sensitive to much of visible spectrum, proper filters are used to obtain maximum tone contrast between healthy and unhealthy foliage at this early date in the development of disease.

Sometimes success in the detection of insects from aerial color infrared photography has been due to the detection of a secondary by-product of the insect presence. Certain types of insects excrete a honeydew substance, a sugary solution that serves as an excellent host medium for the black sooty mold

fungus, *Capnodium Citri* Berk & Desm. This sooty mold fungus coats the leaves of trees and affects the reflection of sunlight from the leaves. Those trees without sooty mold deposits, i.e., without insect infestation, appear bright red, whereas those with infested areas are darker red in color.

The insect-pests and diseases like Brown Soft Scale, Mealy Bugs, Yellow Aphids, Black Scale, Green Scale and White Flies are the examples which produce honeydew which serves as a host medium for a sooty mold fungus. Since the sooty mold deposits blacken the foliage of the trees, these are readily detectable with color infrared photography. Furthermore, due to their different nature in producing honeydew as well as their preference for attacking upper or lower surface of the crown, these insects can be distinguished from one another.

It was hoped that analysis of the color infrared photography and comparison with the ground truth obtained would lead to new insights into the possible causes and treatments of the crop and forest diseases. Although some diseases and insect damage have been detected visually at Weslaco [28-30], in general, the color infrared photography did not have enough contrast for the unaided human eye to differentiate subtle differences in the trees' states of health. Therefore, it is essential to enhance and process the images electronically to obtain the maximum information. The use of electronic enhancement in the studies of a few agriculture diseases such as the corn blight has proved highly successful [31]. Results obtained by Spann [32] et al, show that automatic processing techniques provide an effective tool for studying peach tree decline. At the University of Texas, an interactive system [33-34] has been developed and applied to analyze aerial color infrared photographs.

The research reported in the present paper deals with the detection and classification of infestations caused by the brown soft scale, citrus mealybugs, root rot and gummosis. Brown soft scale produces copious amounts of honeydew, attacks on the upper surface of the crown or leaves, and is capable of squirting the honeydew up to six inches, and hence, deposits laid down by this insect are more continuously distributed over the crown. When they are numerous, the result is like a paint job with a spray. The citrus mealybug, which also produces copious amounts of honeydew, appears to lay down a different pattern of sooty mold deposits than brown soft scale. This is due to the fact that mealy bugs tend to develop inside the crown of the tree and the upper portion of many branches do not become coated with honeydew. It causes the infrared aerial photograph to have a mottled appearance with patches of red color while brown soft scale produces a more continuous appearance. The sooty mold deposits laid down by the citrus blackfly are lighter than those of brown soft scale but they are uniformly distributed over the foliage like the scale, therefore causing differentiable affects on IR film. The trees infected by root rot visually appear healthy even when much of the root system and crown is destroyed by the disease; however, in an IR film the suspected trees appear

lighter red, greenish, and smaller. A tree affected by gummosis appears to be a light pink color in an infrared photograph.

III. MEASUREMENT OF COLOR CHARACTERISTICS

It has been observed that color characteristics of spatial points in the transparency contribute maximum information. However, the qualitative and quantitative measurement and/or comparison of colors is rather elusive. Several methods have been designed to determine color.

In physics, color is related to a specific wavelength of radiant energy. However, this is only one of the three components of the measure of observed color. "Hue" relates to the wavelength of radiant energy which is most predominant in the color. "Purity" describes the absence of other colors, or how close the radiant energy is to a single wavelength. "Intensity" is a measure of the total energy in the color. All three of these parameters are required to describe a color stimuli which is a mixture of the basic primary pure colors (red-green-blue).

One of the first methods developed for the specification of colors was the "Munsell" system [35]. In this technique, the visible spectrum was divided into 50 separate hues, the closest that the human eye can detect. Intensity is measured from 0 (black) to 10 (pure white). Purity is measured from 0 (no color) to a maximum which varies for different hue-intensity values.

The Munsell system was developed for the creation of colors. The measurement of color requires either the comparison of an unknown color with a set of standards in a controlled environment (such as a spectrophotometer), or the detection of the three basic components (red-green-blue). The CIE (Commission Internationale de l'Eclairage) system is based on the measurement of the three basic components and corresponds to the response of the human eye, as reported by the Optical Society of America [36]. The standard observer widely used today was recommended in 1931 by the CIE. The spectral responses of this standard observer are defined by the tristimulus values of the spectrum plotted in Fig. 1. At any wavelength the values of the three functions $R(\lambda)$, $G(\lambda)$, $B(\lambda)$ give the amount of the X (red) primary, the amount of the Y (green) primary, and the amount of the Z (blue) primary, respectively, required to produce that wavelength. The definitions of the tristimulus values x, y, z , corresponding to a stimulus of spectral irradiance, $H(\lambda)$, are:

$$x \equiv \int_0^{\infty} H(\lambda) R(\lambda) d\lambda$$

$$y \equiv \int_0^{\infty} H(\lambda) G(\lambda) d\lambda$$

$$z \equiv \int_0^{\infty} H(\lambda) B(\lambda) d\lambda$$

where $d\lambda$ is an infinitesimal wavelength interval.

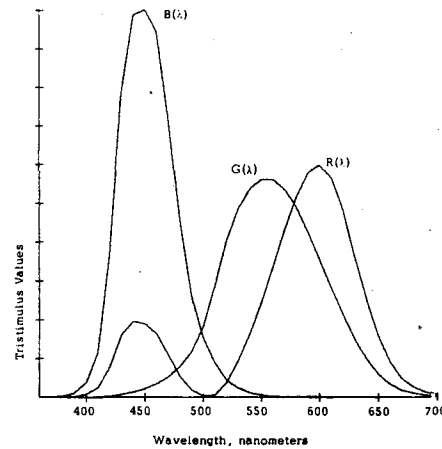


Figure 1. 1931 CIE Standard Observer Color Response

Three quantities called chromaticity coordinates (also called trichromatic coefficients) are defined from the tristimulus values as follows:

$$X = \frac{x}{x + y + z}$$

$$Y = \frac{y}{x + y + z}$$

$$Z = \frac{z}{x + y + z}$$

Since $Z = 1 - X - Y$, X and Y suffice to specify a light source matched by the respective tristimulus values of the three primaries. Therefore, a two dimensional coordinate system, with axes X and Y can be set up so that all light sources are represented by points with the general chromaticity coordinates (X, Y) . These points have been found for all the spectrum colors based on the CIE primaries. The locus of the coordinates for the spectrum colors is called the chromaticity diagram. Fig. 2 shows the positions of common colors in the chromaticity diagram.

A source having equal amounts of radiant power in each wavelength zone of its spectrum would be represented by the chromaticity coordinates $X = 0.333$, $Y = 0.333$ which is called the achromatic point. All the achromatic sources are located in the central region on the chromaticity diagram. If two sources having the chromaticity coordinates (X_1, Y_1) and (X_2, Y_2) are mixed, the resulting mixture is located in the chromaticity diagram somewhere on a straight line joining (X_1, Y_1) and (X_2, Y_2) . If three sources are represented by the points (X_1, Y_1) , (X_2, Y_2) , (X_3, Y_3) , a mixture of the three sources will be somewhere within the triangle having these three points as vertices. Colors within the diagram but outside the triangle cannot be matched by these primaries. This property of the triangle

and the configuration of the chromatic diagram indicate that three spectrum colors - a red, a green, and a blue will match more samples by positive mixing than any other three primaries. Due to this reason, we have chosen these three primaries in our research.

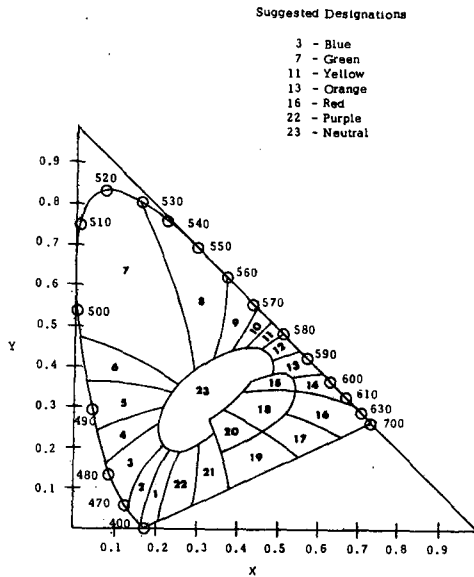


Figure 2. CIE Standard Observer Chromaticity Diagram

IV. DIGITIZATION OF COLOR INFRARED PHOTOGRAPHS

We have developed a system to analyze aerial color infrared photographs of citrus orchards. The input to the system is a 35 mm transparency of the infrared aerial photograph mounted in a 3-color film digitizer. Kodak Ektachrome Infrared Aero Film Type 8443 is used. This film is a false-color, reversal film, sensitive to green, red, and near-infrared radiations instead of the usual blue, green and red sensitization. Healthy green vegetation reflects strongly in the near-infrared region of the spectrum and appears red in the photographs. The film digitizer consists of a high quality flying spot film scanner system which features random access operation and can output a maximum of 8 bits of intensity information with a resolution of 20 lines/mm for each of the red, green and blue channels. The cathode ray tube uses Litton LP203 phosphor, which has two equal energy peaks at 400 and 540 nanometers. Two dichroic filters were used to separate the film colors. The color response of the scanner is shown in Fig. 3.

We have used Ektachrome color infrared film for the detection of plant anomalies because this type of film produces a false color as shown in Fig. 4. The three emulsion layers of the film produce yellow, magenta, and cyan dyes from the green, red, and infrared source reflection. A yellow filter (Wratten No. 12) is used to block the blue color, to which all three layers are sensitive. The

complementary colors of blue, green, and red are formed in the transparency. The major advantage of using this type of film is the red color which is produced by the infrared radiation. The chlorophyll in vegetation is a high reflector of the infrared spectrum. This infrared reflection produces a high contrast of the vegetation with its surrounding area. The degree of redness, hence the amount of chlorophyll in the vegetation, will also determine the type and condition of the vegetation. As shown in Fig. 5, the difference in the measured diffuse reflectance from the heavy and light sooty mold deposits is very little in the visible region (500 to 700 nanometers), but a large difference occurs in the infrared region (700 to 900 nm). This 700 to 900 nm region produces a red color in the Ektachrome infrared film, with the characteristics shown in Fig. 4. Therefore, trees with no sooty mold deposits, hence no insect infestations, appear a bright red color whereas a darker red color is produced as the infestation becomes heavier.

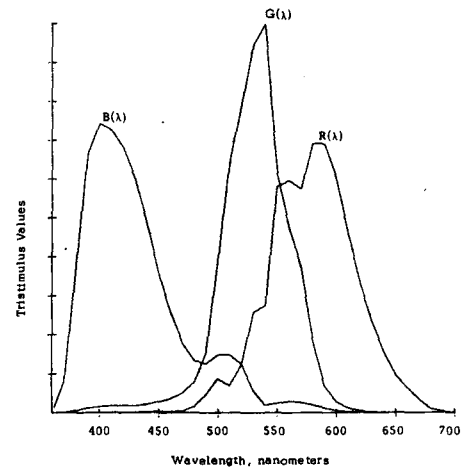


Figure 3. Film Scanner Color Response

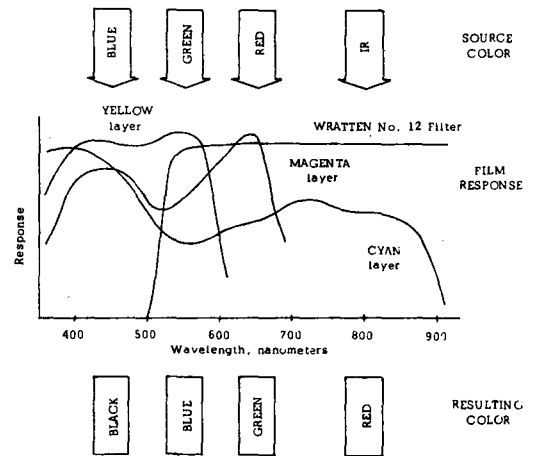


Figure 4. Color Formation by Ektachrome Infrared Film

SOOTY MOLD DEPOSIT
 1-NONE
 2-LIGHT
 3-MEDIUM
 4-HEAVY

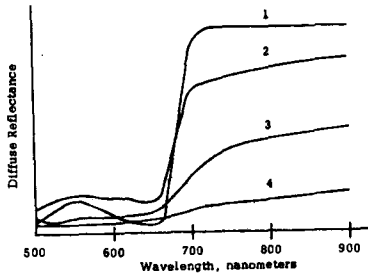


Figure 5. Effect of Sooty Mold on Reflectance from Citrus Leaves

The whole scanner system consists of three main parts: the optics, the electronics, and the software provided by an XDS 930 digital computer which also stores the output information. The optical system projects a spot of light from a precision cathode ray tube onto a 35 mm transparency. As the spot of light is positioned on the film at a preset number of discrete points, the spot is transmitted through the transparency according to the density and color of the film at each location. The resulting light signal is optically changed into its red, green and blue primary components and imaged upon three photomultipliers which transform the light information into electronic signals. Three intensities, of the primaries red, green and blue, of every scanned point of the transparency are transformed into the three normalized color coordinates X, Y, I, which we have used in our research, according to the following equations:

$$X = I_R / (I_R + I_G + I_B)$$

$$Y = I_G / (I_R + I_G + I_B)$$

$$I = (I_R + I_G + I_B) / (3 \times 255)$$

where I_R , I_G and I_B are the measured intensities of red, green and blue channels of the scanner. The gains of the three channels are adjusted such that $0 \leq I_R = I_G = I_B \leq 255$ when a neutral density filter is viewed.

An XDS 930 computer controls the film digitizer, receives the 3-channel color densities in the red, green and blue regions of the spectrum, and generates normalized color coordinates X, Y and I. The film scanner chromaticity diagram, i.e., the loci of normalized color coordinates of common colors, is shown in Fig. 6.

A three dimensional representation of the six

colors is shown in Fig. 7. This figure was produced by our scanner system on a storage display unit as the computer scanned a 35 mm color slide of "Kodak" color control patches. Color purity is based on the assumption that the pure wavelength loci occurs along the boundary of the color triangle.

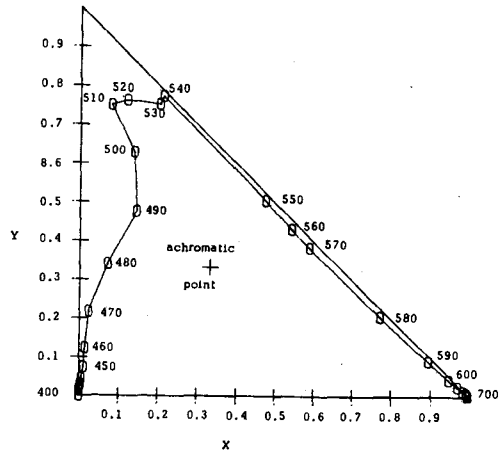


Figure 6. Film Scanner Chromaticity Diagram

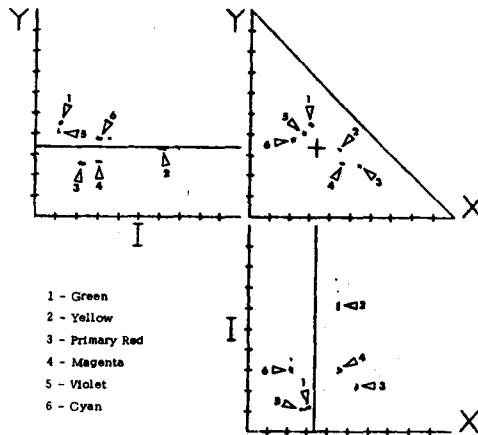


Figure 7. (X, Y, I) for Six Colors

Our system is an interactive system and separation of tree information from the background is performed under operator's direction, on the basis of the difference in color characteristics of the areas having trees and of the background. First a broad color band pass of the three color coordinates (X,Y,I) is chosen which may allow a few background points to pass with the tree points. The computer then displays all the points which can pass through the chosen color band pass. Fig. 8 shows such a display of fifty-nine citrus trees, which resulted by choosing the following broad color band pass:

$$X = 0.60 \pm 0.25$$

$$Y = 0.16 \pm 0.16$$

$$Z = 0.20 \pm 0.20 .$$

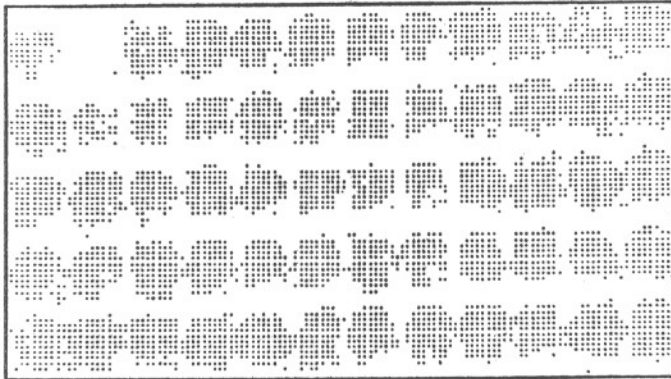


Fig. 8. Computer Display of the Points which are within the Broad Color Band Pass

The difference between the color coordinates of the background and the tree points is measured under operator's direction which refines the color band pass for the sharp separations of the trees from the background. The refined values of the above band pass are as follows:

$$X = 0.58 \pm 0.25$$

$$Y = 0.15 \pm 0.15$$

$$I = 0.23 \pm 0.23 .$$

A three dimensional representation of all the color coordinates of the scene and the refined color band pass is shown in Fig. 9.

The computer then considers only those points of the scene which are within the refined color band pass, locates the outline of each individual tree, and displays all the tree points with the tree boundaries. Such a display of fifty-nine trees of Fig. 8 is shown in Fig. 10.

V. CLUSTERING OF TREES ACCORDING TO INFESTATIONS

Further analysis is carried out on an individual tree basis, i.e., the algorithm considers the points inside the boundary of each individual tree separately. Cluster analysis is applied by using the normalized color coordinates of the points of each individual tree. Some methods have been put forward by different authors for the clustering of individuals. A review and a large number of references are given by Cormack [37] and Kendall [38]. However, they do not by any means lead to identical or even to similar results. We have derived an empirical distance function to cluster the trees on the basis of the type and amount of infestation

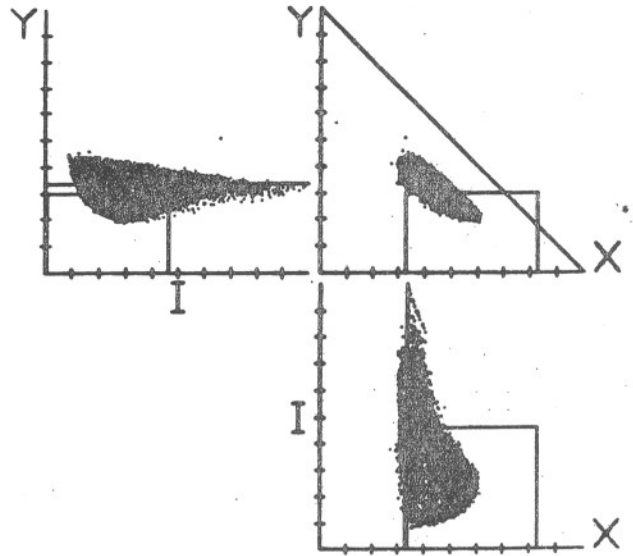


Fig. 9. A Three Dimensional Representation of all the Color Coordinates of the Scene and Color Band Pass

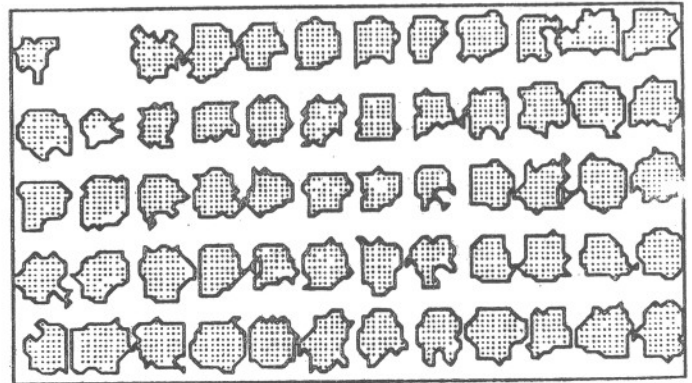


Fig. 10. Computer Display of Fifty-nine Trees

present. Let us assume there are n trees in a transparency which are to be clustered on the basis of p variables. For a set of n individuals there are $\frac{1}{2}n(n-1)$ distances between pairs. To calculate and consider them all is computationally prohibitive if n is large, e.g., for $n = 1000$ there are nearly half a million distances. On the other hand, one can compute the center of gravity (C.G.) and the distances of the n points from it. The computation is then proportional to n , not to n^2 . We have, therefore, followed the latter approach in our algorithm.

We have used six variables (i.e., $p = 6$) which

are the means and standard deviations of the three normalized color coordinates X, Y, I, of the points of a tree. These six variables v_{ij} ($i = 1, 6$) are represented as follows for the j -th tree having m_j points:

$$v_{1j} = \frac{1}{m_j} \sum_{k=1}^{m_j} X_k$$

$$v_{2j} = \frac{1}{m_j} \sum_{k=1}^{m_j} Y_k$$

$$v_{3j} = \frac{1}{m_j} \sum_{k=1}^{m_j} I_k$$

$$v_{4j} = \sqrt{\frac{\sum_{k=1}^{m_j} (X_k - v_{1j})^2}{m_j - 1}}$$

$$v_{5j} = \sqrt{\frac{\sum_{k=1}^{m_j} (Y_k - v_{2j})^2}{m_j - 1}}$$

$$v_{6j} = \sqrt{\frac{\sum_{k=1}^{m_j} (I_k - v_{3j})^2}{m_j - 1}}$$

The Center of Gravity (C.G.) of the set of n trees of a given transparency is represented by the means of the first three variables, that is:

$$\text{C.G.} = (\bar{v}_1, \bar{v}_2, \bar{v}_3)$$

where $\bar{v}_i = \frac{1}{n} \sum_{j=1}^n v_{ij}$ for $i = 1, 3$.

The distance function, we have empirically derived, computes the tree distances from the Center of Gravity. The tree distance, TD_j , of the j -th tree from the C.G. is represented as follows:

$$\begin{aligned} TD_j &= w_1 \cdot D_{1j} \cdot S_{1j} \cdot N_{1j} \\ &+ w_2 \cdot D_{2j} \cdot S_{2j} \cdot N_{2j} \\ &+ w_3 \cdot D_{3j} \cdot S_{3j} \cdot N_{1j} \end{aligned}$$

where

$$D_{ij} = (v_{ij} - \bar{v}_i)^2$$

$$S_{ij} = \frac{v_{(i+3)j} \cdot |v_{ij} - \bar{v}_i|}{A_{ij}}$$

$$A_{ij} = \frac{1}{n} \sum_{j=1}^n |v_{ij} - \bar{v}_i|$$

for $i = 1, 3$

$$N_{ij} = \frac{(v_{ij} - \bar{v}_i)}{|(v_{ij} - \bar{v}_i)|}$$

for $i = 1, 2$.

w_1 , w_2 , and w_3 are weighting variables determined empirically. The values of these variables, determined on the basis of the analysis of the data (collected from a number of photographs) are: $w_1 = 5.0$, $w_2 = 2.0$ and $w_3 = 0.5$.

D_{ij} is the square of the distance of the j -th tree from the center of gravity on the i -th coordinate axis ($i = 1, 3$). The standard deviation of each coordinate, within the j -th tree is divided by the average of the absolute values of the distances of the trees of the transparency from the center of gravity and is multiplied by the absolute value of the distance of the j -th tree from the center of gravity. The standardized form of the standard deviation is represented by S_{ij} . Such a standardization was necessary to coordinate the variation of the standard deviation with the variation of the distance from C.G. It has been observed that although standard deviation contributes a lot in resolving the healthy and infested trees, its unstandardized form sometimes introduces confusion. For example, light infestations scattered on the crown of the tree cause a large standard deviation in comparison to healthy trees which have uniform distribution of color. This enlarges the usually positive tree distance of the lightly infested tree in comparison to the healthy tree. Similar confusion arises between the scattered medium infestations and heavy infestations covering the entire crown of the tree.

The trees having positive tree distances lie on one side of C.G. and those having negative tree distances lie on the other side of C.G. The sign of the tree distance TD_j depends on the amplitudes and signs of its three components. The signs of these components depend on N_{ij} which is either +1 or -1 depending upon the sign of the distance of j -th tree from C.G. on the i -th coordinate axis (for $i = 1, 2$). Since the boundaries on the third coordinate axis (i.e., I) are not very sharp, we have used N_{1j} as the sign factor on that axis.

It has been observed that healthy trees having maximum positive tree distances from C.G. lie at one end and heavily infested trees or bushes of grass (which sometimes grow at the place of dead trees) having maximum negative tree distances lie at the other end. The trees with light and medium infestations fall in between these two limits in such a way that the trees with light infestations are closer to healthy trees and the trees with medium infestations are closer to the trees having

heavy infestations. Depending upon the type and amount of infestation and the number of infested trees, the center of gravity may itself lie in any of these regions. However, generally it falls in the region of medium infestations.

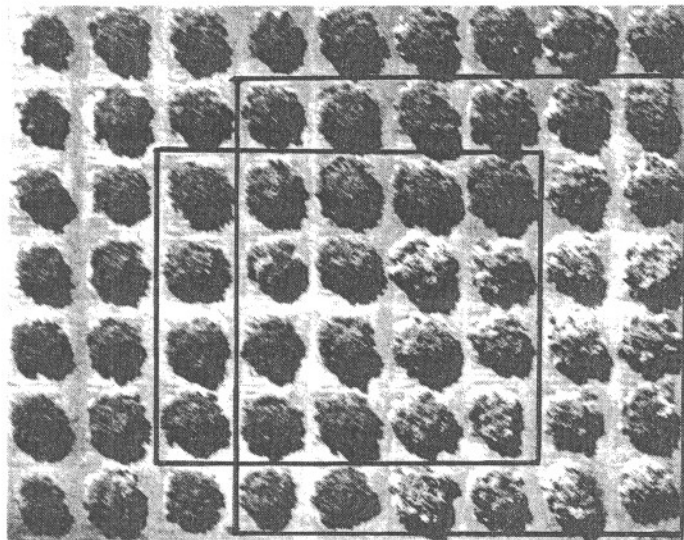


Fig. 11. Citrus Trees having Gummosis Infestation

Sorting, performed on the basis of the computed tree distances, clusters the trees. However, the process of cluster identification cannot be accomplished by simply setting the thresholds. Table I shows data computed from the measurements of the transparency, shown in Fig. 11, in which some of the trees have Gummosis infestation. Thirty-six trees of the photograph, boxed in Fig. 11, are considered. Plots of this data are shown in Fig. 12. The tree numbers in the first column have been sequenced after sorting the trees on the basis of tree distances. Sorted tree distances are shown in the second column. Since the tree distances at the top and at the bottom of this column are large, their differences are also supposed to be large in comparison to the smaller tree distances from C.G. which are in the middle of the column. A plot of these tree distances is shown in Fig. 12A (X-axis is in the middle). The differences of the tree distances are plotted in Fig. 12B (X-axis is at the bottom) and depicted in the third column of Table I.

The difference in TD_j 's of the two adjacent trees, which are also at the boundary of two adjacent clusters, should be larger than the difference of TD_j 's of those adjacent trees which are within the clusters. Hence, the peaks in the graph of Fig. 12B represent the cluster boundaries. These cluster boundaries are also depicted in the fourth column of Table I.

Since the clusters are formed on the basis of color characteristics of the trees, many times the number of clusters may be larger (e.g., fifteen clusters are formed in the example shown in the

TABLE I: Computed Tree Measurements

TREE NO.	TREE DISTANCES	TREE DIST. DIF.	CLUST. BOUNDARIES	THRESHOLDS
1	131730	3831		
2	127899	47402	47402	
3	80497	18029		
4	62468	5427		
5	57041	11712	11712	
6	45329	1160		
7	44169	1825	1825	
8	42344	1208		
9	41136	5018	5018	
10	36118	4848		
11	31270	9331	9331	9331
12	21939	239		
13	21700	4180	4180	
14	17520	543		
15	16977	4475		
16	12502	5420	5420	5420
17	7082	750		
18	6332	2305	2305	
19	4027	678		
20	3349	1023		
21	2326	2266	2266	
22	60	919		
23	-859	2848	2848	
24	-3707	999		
25	-4706	18254	18254	
26	-22960	9910		
27	-32870	3570		
28	-36440	22381	22381	22381
29	-58821	75		
30	-58896	9219	9219	
31	-68115	7332		
32	-75447	18870		
33	-94317	20136		
34	-114453	37528	37528	
35	-151981	28664		
36	-180645			

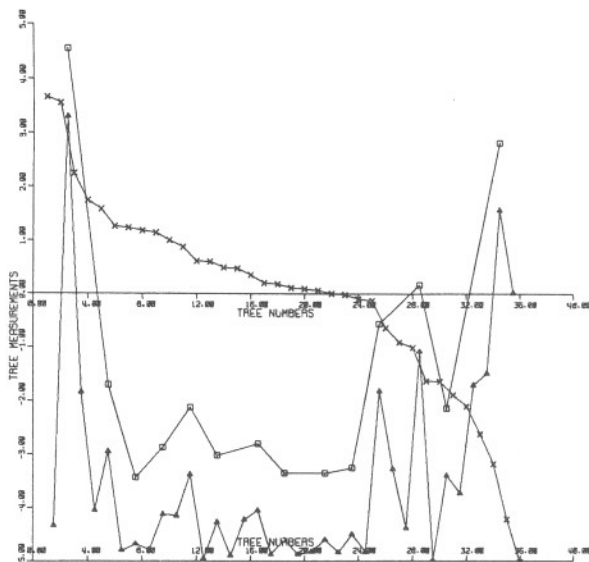


Fig. 12. Plot of (A) Tree Distances [x-x] (B) Dif. of Distances [$\Delta - \Delta$] (C) Cluster Boundaries of 36 Trees [$\square - \square$]

fourth column of Table I) than the conventionally understood four or five categories like heavy, medium, light and no infestations. Therefore, the clusters are to be combined to form fewer groups. The process of cluster identification, described in the previous paragraph, can be repeated to separate out the new groups. The plot of cluster boundaries is shown in Fig. 12C and the peak positions of this

graph represent the separating thresholds of the new groups. These thresholds are also shown in the fifth column of Table I. In this way fifteen clusters are grouped into four classes. The trees of the first and last groups possess the healthy and heavily infested trees, respectively. The second and third groups possess the trees having light and medium infestations, respectively.

Such a process of cluster identification can be repeated till the required number of groups is obtained. However, the number of groups which exist in a transparency may not be exactly equal to the desired number of groups. In that case the groups in between the first and last group can be graded from light to medium infestations, depending on the number of these groups. Sometimes the desired number of groups result even in the first attempt.

VI. EXPERIMENTAL RESULTS

We have analyzed about fifteen infrared aerial color photographs of citrus orchards having infestations of Citrus Mealybugs, Brown Soft Scale, Gummosis and Root Rot. Figure 11 depicts the photograph of citrus trees having Gummosis infestations. Two samples of twenty trees and thirty-six trees of this photograph have been analyzed separately. These two groups of trees are marked separately in Fig. 11. The computer display of the tree boundaries of the set of twenty trees is shown in Fig. 13.

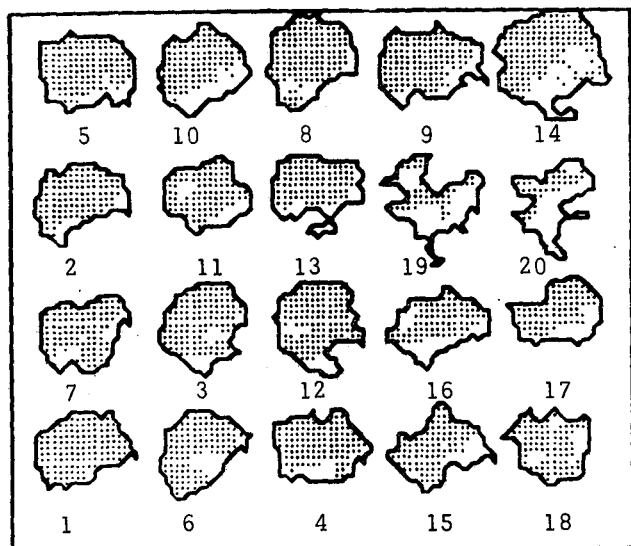


Fig. 13. Computer Display of Twenty Trees of Fig. 11

The spatial positions of the trees in Fig. 13 correspond to the spatial positions of the trees marked in Fig. 11; however, the illustrated tree numbers are in the sequence of sorted trees. This sequence of trees is shown in the first column of Table II. The sorted tree distances, their differences and cluster boundaries are shown in the second, third and fourth columns of this table. The plots of these measurements are shown in Fig. 14. Since only four clusters are formed in the first attempt, no further grouping is needed in this case and hence the cluster boundaries also represent the

group thresholds. The tree measurements and their plots, shown in Table I and Fig. 12 respectively, of the second group of thirty-six trees have already been explained in the previous section, i.e., fifteen clusters are formed which are grouped into four classes. (The words "groups" and "classes" have been used in the same sense here.)

TABLE II: Computed Measurements of Twenty Trees of Fig. 11

TREE NO.	TREE DISTANCES	TREE DIST. DIF.	CLUST. BOUNDARIES OR THRESHOLDS
1	354884	283834	
2	71050	26719	
3	44331	1811	
4	42520	19691	19691
5	22829	7950	
6	14879	1337	
7	13542	2709	
8	10833	6659	6659
9	4174	2237	
10	1937	415	
11	1522	332	
12	1190	920	
13	270	1133	
14	-863	4425	
15	-5288	6959	
16	-12247	19863	
17	-32110	45745	
18	-77855	409505	409505
19	-487360	239394	
20	-726754		

The two sets of the trees of the same transparency were considered here for the convenience of depicting the data as well as illustrating different aspects of the algorithms. For the former reason the number of trees in these samples was also kept small. Generally, in practice, the number of trees analyzed together is large. Our experimental results show very satisfactory classification of infestations and agree with the human interpretations as well as with the data collected through the ground survey.

VII. CONCLUSION

The algorithms and cluster analysis used in this paper are fairly insensitive to color variations of the transparencies due to various photographic and environmental effects. This is an important quality of the system because the color characteristics of a transparency depend on the weather conditions at the time of photography. For example, a photograph taken in sunshine results in a lighter transparency than one taken in cloudy conditions. In addition, different varieties of trees produce different color effects on the transparency, e.g., among citrus trees, orange, grapefruit and tangerine produce different color characteristics. Therefore, if any system is sensitive to any of these types of color variations besides the color variations caused due to the infestations, it will affect the accuracy and reliability of the classification.

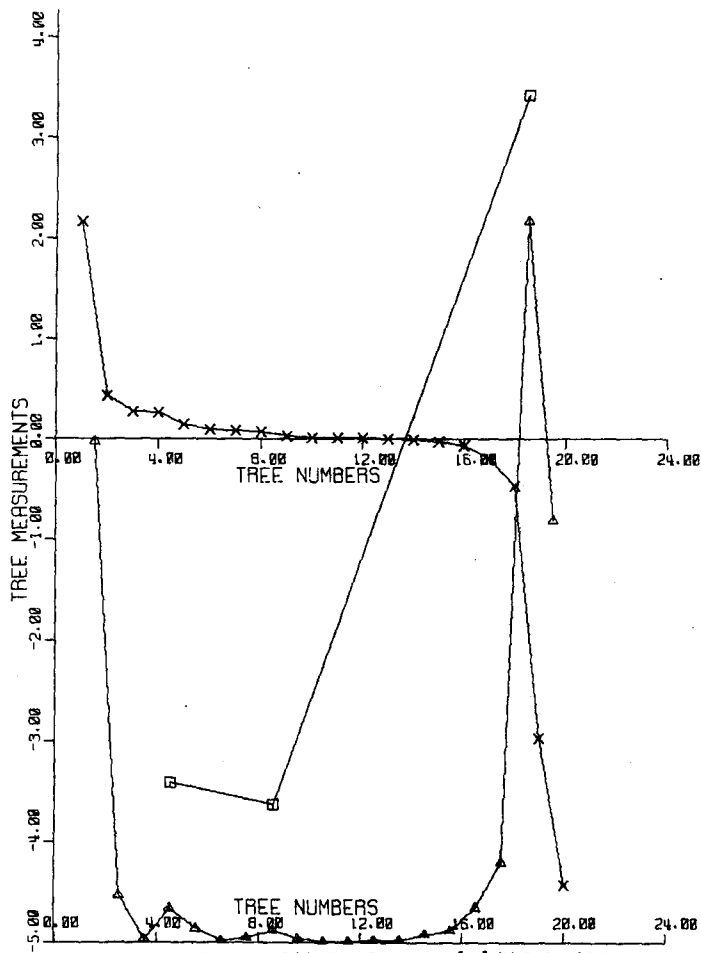


Fig. 14. Plot of (A) Tree Distances [x-] (B) Dif. of Distances [△] (C) Cluster Boundaries of 20 Trees [□]

It has been observed that different types of infestations produce different types of color variations in the infrared aerial color photographs. Hence it is possible to classify crop insect pests and diseases on the basis of their effect in changing the color characteristics in the transparency. We have been working on this type of classification; the results of our finding will be reported in our next publication.

ACKNOWLEDGMENTS

The authors wish to thank Mr. W.G. Hart, Mr. R. Davis, and Mr. S.J. Ingle, of the United States Department of Agriculture, Agricultural Research Division, Entomology Research Division, Weslaco, Texas, for the aerial color infrared photographs, and for their valuable assistance in the human interpretation of the photographs. We also wish to thank Mr. D. Williams and Mr. M. Day for their assistance with the color film scanner.

The research reported herein was supported by NSF Grant GK 42790, by AFOSR Contract 72-2371, and by JSEP Program under Contract F44620-70-C-0091.

BIBLIOGRAPHY

1. U.S. Agricultural Research Service, 1965. "Losses in Agriculture," Handbook No. 291. Dept. Agr., Washington, D.C., pp. 3-54.
2. Prince, V.E., 1966. "Winter injury to peach trees in Central Georgia," Proc. Amer. Soc. Hort. Sci. 88: 190-196.
3. Clayton, C.N., 1971. "The perennial (Valsa/cytospora) canker complex," Proc. of National Peach Council 30th Ann. Convention, 33-34.
4. Cowart, F.F. and E.F. Savage, 1941. "Important factors affecting peach tree longevity in Georgia," Proc. Amer. Soc. Hort. Sci. 39: 173-176.
5. Dowler, W.M. and D.H. Petersen, 1966. "Introduction of bacterial canker of peach in the field," Phytopathology, 56: 989-990.
6. Hendrix, F.F., Jr., W.M. Powell, and J.H. Owen, 1966. "Relation of root necrosis caused by pythium species to peach tree decline," Phytopathology, 56: 1229-1232.
7. Higgins, B.B., C.P. Walton, and J.J. Skinner, 1948. "The effect of nitrogen fertilization on cold injury of peach trees," GA. Agric. Exp. Sta. Bull. No. 226, 27 pp.
8. Hine, R.B., 1961. "The role of fungi in the peach replant problem," Pl. Dis. Repr. 45: 462-465.
9. Luepchen, N.S. and K.G. Rohrbach, 1969. "Cytospora canker of peach trees: spore availability and wound susceptibility," Pl. Dis. Repr. 53: 869-872.
10. Petersen, D.H. and W.M. Dowler, 1965. "Bacterial canker of stone fruits in the southeastern states," Pl. Dis. Repr. 59: 701-702.
11. Prince, V.E., 1966. "Winter injury to peach trees in Central Georgia," Proc. Amer. Soc. Hort. Sci. 88: 190-196.
12. Savage, E.F. and F.F. Cowart, 1942. "Factors affecting peach tree longevity in Georgia," Ga. Agr. Expt. Sta. Res. Bull. 219, 15 pp.
13. Savage, E.F. and F.F. Cowart, 1954. "Factors affecting peach tree longevity in Georgia," Proc. Amer. Soc. Hort. Sci., 64: 81-86.
14. Savage, E.F. and R.A. Hayden, and W.E. Ward, 1963. "The effect of type and season of pruning on growth and yield of Dixigen peach trees," Ga. Agr. Exp. Sta. Bull. N.S. 109, 18 pp.
15. Savage, E.F., J.H. Weinberger, E.S. Luttrell, and A.S. Rhoads, 1953. "Clitocybe root rot - a disease of economic importance in Georgia peach orchards," Pl. Dis. Repr. 37: 269-270.
16. Taylor, J., J.A. Hiesbrock, F.F. Hendrix, W.M. Powell, J.W. Daniell, and F.L. Crosby, 1970. "Peach tree decline in Georgia," Ga. Agr. Expt. Sta. Res. Bull. 77, 45 pp.
17. Newell, W., and A.C. Brown, 1939. "Eradication of the citrus blackfly in Key West, Florida," J. Econ. Entomol. 32: 680-2.
18. Smith, H.D., and H.L. Maltby, 1964. "Biological Control of the citrus blackfly in Mexico," USDA Tech. Bull. 1311, 30 pp.
19. Payne, Jerry A., W.G. Hart, M.R. Davis, L.S. Jones, D.J. Weaver, and B.D. Horton, 1971. "Detection of peach and pecan pests and diseases with color infrared aerial photography," Proceedings of the Third Biennial Workshop on Color Aerial Photography in the Plant Sciences, Gainesville, Florida, March 2-4.
20. Turner, W.F., 1949. "Insect vectors of phony peach disease," Science 109 (2822): 87-88.
21. Turner, W.F. and H.N. Pollard, 1955. "Additional leafhopper vectors of phony peach," J. Econ. Entomol. 48(6): 771-2.
22. Pritchard, A.E. and E.W. Baker, 1952. "A guide to the spider mites of deciduous fruit trees," Hilgardia 21(9): 253-87.
23. Payne, J.A., W.L. Tedders, and C.R. Centry, 1970. "Pecan borers threaten Georgia's crop yields," The Pecan Quarterly 4(1): 3-4.
24. Cole, J.R., 1937. "Bunch disease of pecans," Phytopathology 27(5): 604-12.
25. Colwell, R.N., 1956. "Determining the prevalence of certain cereal crop disease by means of aerial photography," Hilgardia, 26(5): 223-286.
26. Colwell, R.N., 1965. "The extraction of data from aerial photographs by human and mechanical means," Photogrammetria, 20: 211-228.
27. Langley, P.G., 1965. "Automatic aerial photo-interpretation in forestry - how it works and what it will do for you," Proc. Ann. Meeting Soc. Amer. For., Detroit, Michigan.
28. Hart, W.G. and V.I. Myers, "Infrared aerial color photography for detection of populations of brown soft scale in citrus groves," J. Econ. Entomol. vol. 61, no. 3, June 1968, pp. 617-624.
29. Hart, W.G., S.J. Ingle, R. Davis, C. Mangum, A. Higgins, and J.C. Boling, "Some use of infrared aerial color photography in entomology," Proceedings of the Third Biennial Workshop on Color Aerial Photography in the Plant Sciences, March 1971.
30. Hart, W.G., S.J. Ingle, M.R. Davis, and C. Mangum, "Aerial photography with infrared color film as a method of surveying for citrus blackfly," J. Econ. Entomol., vol. 66, no. 1, Feb. 1973, pp. 190-194.
31. MacDonald, R.B., M.E. Bauer, R.D. Allen, J.W. Clifton, J.D. Erickson, and D.A. Landgrebe, 1972. "Results of the 1971 corn blight watch experiment," Proc. of the Eighth International Symposium on Remote Sensing of Environment, Vol. 1, pp. 157-189.
32. Spann, G. William, J.A. Payne, and H.C. Kirkpatrick, "Infrared photography of peach short life sites in Georgia," Proceedings of the Fourth Biennial Workshop on Color Photography in the Plant Sciences, Orono, Maine, July 10-12, 1973.
33. Underwood, S.A. and J.K. Aggarwal, "Computer Analysis of Aerial Color Photographs," Technical paper accepted in Journees d'electronique 1975 on Advanced Processing Technology, Switzerland.
34. Underwood, S.A., D.H. Williams, P. Costa, and J.K. Aggarwal, "A flying spot scanner system," Technical Report No. 164, Information System Research Laboratory, Electronics Research Center, The University of Texas, Austin, Texas, Dec. 1974. To appear in the Transactions of the Instrument Society of America, Vol. 14, No. 4, April 1976.
35. Munsell, "A Grammar of Color," edited by Fager Birren, Van Nostrand Reinhold Company, 1969.
36. "The Science of Color," Committee on Colorimetry, Optical Society of America, Thomas Y. Crowell Company, 1953.
37. Cormack, R.M., 1971. "A review of classification," J. Roy. Statist. Soc., 134, 321-367.
38. Kendall, M.G., 1971. "Cluster Analysis," Proceedings Hawaiian Conference on Pattern Recognition.