

AGRICULTURE DISEASE MITIGATION SYSTEM

Sion Hannuna¹, N. Anantrasirichai², Swarna Subramanian³, Suma Prashant⁴, Ashok Jhunjunwala⁵ and C. Nishan Canagarajah⁶

¹Department of Computer Science, University of Bristol, UK
E-mail: hannuna@bristol.ac.uk

^{2,6}Department of Electrical and Electronic Engineering, University of Bristol, UK
E-mail: ²N.Anantrasirichai@bristol.ac.uk, ⁶Nishan.Canagarajah@bristol.ac.uk

^{3,4}IITM's Rural Technology and Business Incubator, India
E-mail: ³swarna@rtbi.in, ⁴suma@rtbi.in

⁵Department of Electrical Engineering, Indian Institute of Technology Madras, India
E-mail: ashok@tenet.res.in

Abstract

Around 52% of the population of India rely on farming for their livelihood which accounts for 17% of India's GDP. Whilst most farmers are familiar with conventional farming practices, they are often ill positioned to promptly deal with diseases and plant infestations affecting their crops. Current advisory systems tend to be generic and are not tailored to specific plots or farms. This work comprises an agriculture advisory call center similar to a modern call center to provide an agriculture disease mitigation system. The information regarding an individual farm is collected using mobile phones. The image of diseased/infected crop is also captured using mobile phones and is made available to the expert to provide the advisory. To scale the advisory, an attempt is also made to automate the disease recognition process using image processing. Unfortunately, the photos taken will be sensitive to a number of factors including camera type and lighting incident on the scene. Ideally, the images would be processed in such a way as to provide the expert with a visual representation of the affected crops that reflects the true nature of the scene. We describe a framework for standardising the colour of plant images taken using both mobile phones and compact cameras within the context of the advisory system.

Keywords:

Plants, Pathology, Colour, Characterisation, Agriculture Advisory, Disease Mitigation

1. INTRODUCTION

At the time of independence, most of India used traditional agricultural practices, where the knowledge was passed from generation to generation, seeds were grown locally, local organic pesticides and fertilizers were used and the market was largely local. Farmers had control over all aspects except the weather. With the increasing demand for food to feed the growing population, India switched to what is known as modern scientific agriculture [1] somewhere between 1960 and 1980; chemical fertilizers and pesticides were now used, seeds came from outside the village and the markets gradually became global. The knowledge itself was now in hands of agricultural scientists and manufacturers of seeds and fertilizers. The cost of inputs became high increasing the risks of failures. Weather risk, risk due to diseases and market-related risks as the global prices fluctuated made agriculture a highly risky business.

1.1 INFORMATION GAP

This changing scenario created a large information gap [2] and farmers depended on outsiders (agricultural scientists, extension workers, vendors of agricultural inputs) for knowledge and experts for markets. The farmers lost control of most farming practices, the predominant one being disease mitigation. When a pest or a disease strikes a farm, it requires an instant response; two to three days would be too late. The Government had created an agricultural extension system, to act as a bridge between the agricultural scientists, researchers and the farmer. The problems faced by the farmers are to be presented to the agricultural scientists for solutions; similarly, the research findings are to be implemented in the field for farmers' benefit. However, the extension system fails to reach small and medium farmers. Fig.1 shows the extent of support that different class of farmers receive from different sources [3]. Further the extension system is slow and does not respond fast enough for disease mitigation; while some of the common pests and disease can indeed be handled by the general information provided by different sources, pests and disease specific to individual farms just cannot be handled by such support systems. The result is a huge loss to farmers; the losses of agricultural produce due to pests alone amounts to Rs. 1.40 Lakh Crores or 17% of total agricultural produce [4].

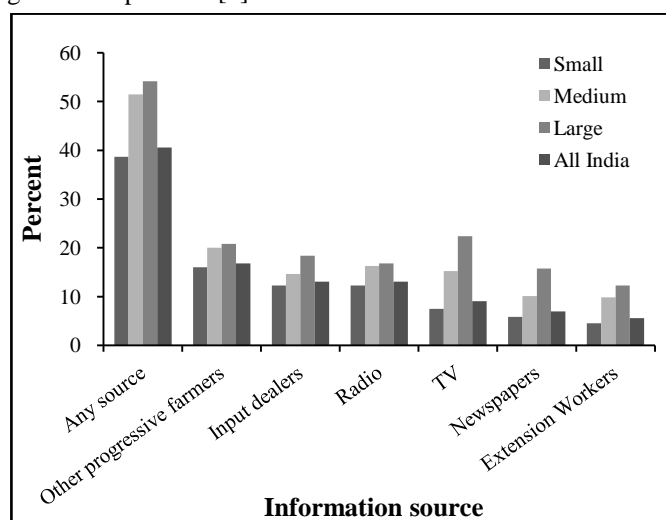


Fig.1. Information sourced on modern technology by farmers according to landholding size, from 2003 NSSO survey

1.2 ABOUT THIS PAPER

This paper presents an effort of using ICT to overcome the information gap by:

- Creating a sophisticated call center akin to what is used today in the developed world to provide support services for goods and merchandise sold to individual customers.
- Attempting to automate as far as possible disease recognition process using image processing.

2. MODERN CALL CENTER

The modern call center uses sophisticated technologies. First of all an extensive data is captured for each customer including specifics of the purchase, the past interaction that he/she would have had with the call center and the complete service history. As the customer calls, the call is routed to an agent and the customer-page pops up on the agent's screen. Within seconds, the agent gets a picture of the profile of the customer, the service history making the interaction personalized. Further, the interaction is continuously captured in the database and a supervisor is added to the conference call whenever the customer is not satisfied with the present interaction [5].

2.1 AGRICULTURE ADVISORY CALL CENTER

The agriculture advisory call center being built is similar to that used in a modern call center and has the following features:

- Uses state-of-the-art technology
- Has a database of the farmer and farm
- The database is constantly updated
- Provides all the previous interactions with the farmer.
- Integrates relevant information (like recent weather-history at the farmer's land, prices at nearby markets) and presents them on a dashboard.
- Makes images of diseased crop available to assist in providing the advisory.
- Gives the extension worker at the call center a virtual feel of being at the farm.
- On an as needed basis scientific and market experts are added to the conference call to provide specific advisory.

The agriculture call center has an information repository of the individual farms collected using a mobile phone for data and image capture. Continuous updating of data is carried out and the interactions with call center advisory are captured. Farmers can upload additional information and pictures of the crop using their mobile phones. The extension worker handles the calls from farmers and conferences in experts as and when necessary.

3. COMPUTER BASED DISEASE RECOGNITION PROCESS

When a farmer calls the call center for advice, the relevant information of the farm is presented on a dashboard. When needed, the farmer is requested to upload the pictures of diseased crop. The extension worker uses the crop information along with its images to provide the advisory. When the extension

personnel fails to identify the disease or the pest or do not have the solution, an expert or a scientist is added to the conference call. The experience is that the solution is found quickly most of the times.

While the system does work, it would be difficult to scale the system as more and more farmers start using it. Trained extension personnel as well as agricultural experts / scientists are limited in number. Using their services to handle each call would not be scalable. The next step would be to automate the disease/pest detection and advisory system. The paper presents some early work towards automatic disease/pest detection based on images sent by the farmers. As the advisory is based on the image, the quality of the image becomes a critical factor. The automatic disease recognition process incorporates the following:

- Colour Correction
- Standardizing for Lighting and Camera Variation
- Deducing a Transformation Matrix
- Leaf Segmentation
- Disease Recognition

Fig.2 depicts the process flow of the Agriculture Disease Mitigation System with the Disease Recognition Process.

3.1 COLOUR CORRECTION

The appearance of images depends on various elements. On the acquisition side, it is influenced by the interaction of the scene's content, the lighting conditions and the capture device. On the display side it is affected by the screen set-up and its local environment with both influencing how the human visual system perceives the scene [6]. For example the same image will look different under artificial and natural lighting. Note that these perceptual differences are not just a by-product of the HVS: the actual pixel values of the same object are highly dependent on capture conditions. This makes the use of colour and intensity information as features for automatic analysis very problematic.

In this paper we propose a cheap and easily implemented approximation to such a solution that makes use of automatically detected colour charts facilitating its integration into colour-managed imaging pipelines. The motivation for maximising how realistic the plant images are is that a great deal of information about the plants' health may be gleaned from their colours. Fig.3 shows two red maple leaves afflicted with different mineral deficiencies. Different capture conditions can greatly affect how those images appear. In addition to striving for perceptual consistency, objective colour consistency increases the potential for fully exploiting colour as a feature in automated systems.

3.2 STANDARDIZING FOR LIGHTING AND CAMERA VARIATION

In order to obtain an independent measure of the colour in an image, colour targets are placed in the scene which is to be captured. A common target widely used in colorimetry and photography is the X-Rite (former Gretag-MacBeth) ColorChecker. In our trials we tested both this target and one of our own devising that has been tailored to our needs. Specifically, a white border has been added to facilitate

automatic detection and ‘greenish’ colours have been more numerously represented. It was anticipated that by more heavily weighting areas of the gamut where foliage colours generally reside, they may be more accurately transformed.

As this work involves deployment in rural India, the replication cost of the colour targets has to be low. This is

achieved by using a traditional chemical printing process creating costs of less than 10 cents per target. To avoid specular highlights in the colour target, it is printed on matte photo-paper. A further design decision was to print the target in credit-card size. This enables the farmer to have the target at hand when needed.

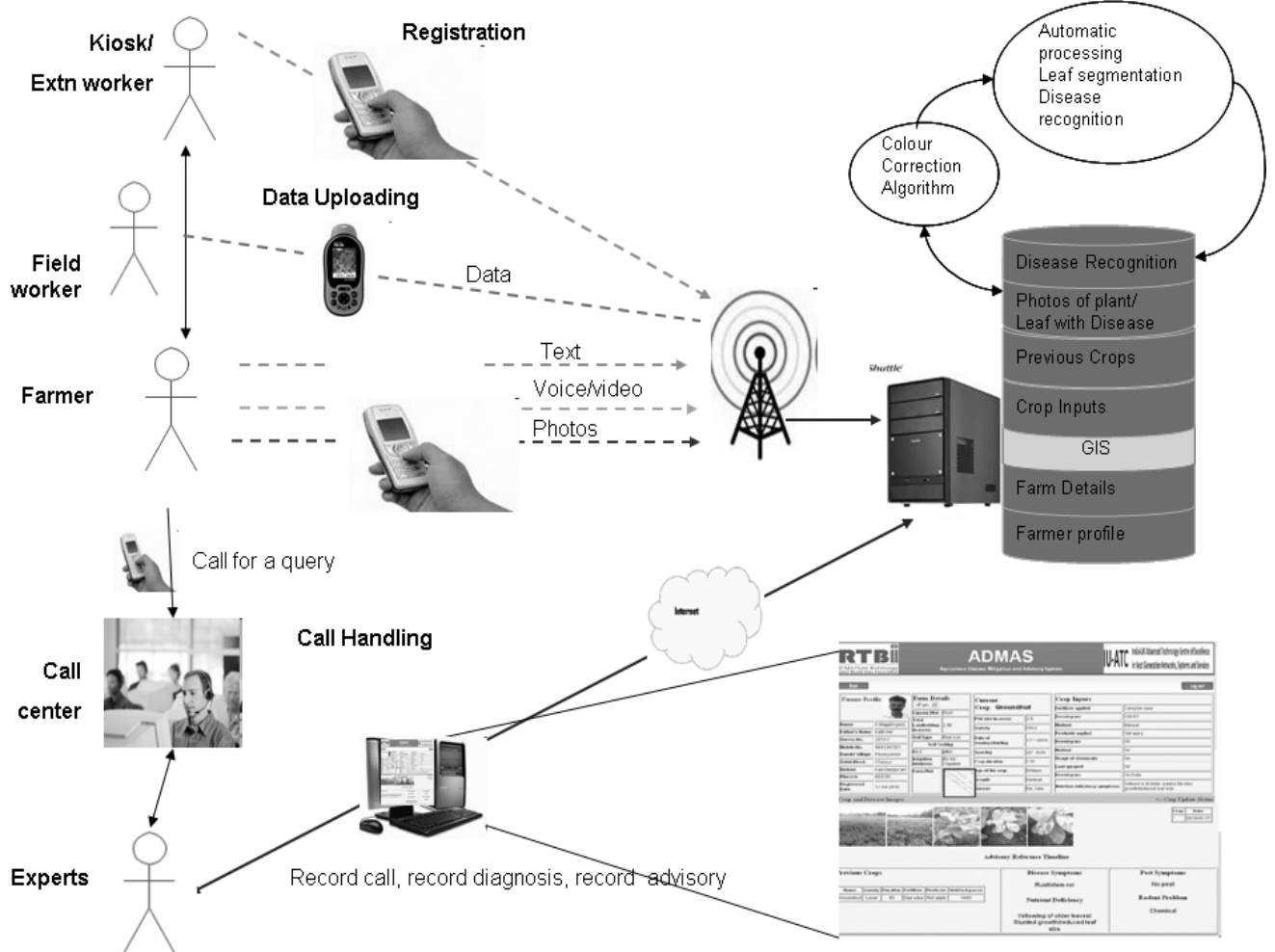


Fig.2. Process Flow Diagram



Fig.3. Iron (a) and magnesium (b) Deficiency in red maple (Acer rubrum L) Images by John Ruter, University of Georgia, Bugwood.org

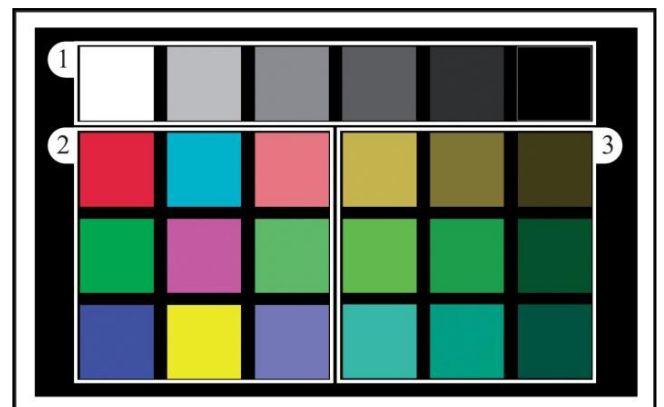


Fig.4. Custom colour target designed to preferentially map plant colours and enable automatic segmentation. Note that the white numbered boxes in the Fig. are for illustration purposes only and not part of our implemented target

3.3 DEDUCING A TRANSFORMATION MATRIX

The aim of our system is to standardise images from low quality cameras such that they are both consistent with one another and provide a faithful representation of the scene they depict.

The patch values for the source image to be processed and the ground truth image are used to deduce a transformation for the source image. This process is analogous to the characterisation of imaging devices including digital cameras. Various characterisation methods exist [7], [8]. They essentially involve deducing a mapping for targets that have known device independent CIE XYZ values. Linear and polynomial transformations have been used to this end as well as neural networks. However, an important distinction for our application should be noted: normally characterisation is done *once for a given device*. Hence it is important that the mapping used generalises for all conceivable scenes and lighting conditions. Conversely, in our application, we only need the mapping to work well for one particular image.

To exploit this stipulation we have experimented with weighting the relative cost for the residual error of particular patches under the deduced transformation according to how numerous image pixels closest to that patch are in a given image. Hence, in most of our images greens and browns would be favoured as they primarily depict plants. For the quadratic and linear transformations we utilise, this lends itself to a weighted least squares optimisation. We term this our ‘heuristic’ approach, whereas normal least squares will hereafter be referred to as ‘pure’.

The colour targets we tested only have 24 patches. We found this was not enough correspondences for training MLP neural networks and Support Vector Machines as generalisation was poor. Conversely, the quadratic and linear transformations considered here produce perceptually convincing results. Cheung and Westland describe a quadratic mapping function [9] formulation that may be solved as a linear system

$$\begin{aligned}
 X &= a_1R + a_2G + a_3B \\
 &+ a_4RG + a_5RB + a_6GB \\
 &+ a_7R^2 + a_8G^2 + a_9B^2 + a_{10} \\
 Y &= a_{11}R + a_{12}G + a_{13}B \\
 &+ a_{14}RG + a_{15}RB + a_{16}GB \\
 &+ a_{17}R^2 + a_{18}G^2 + a_{19}B^2 + a_{20} \\
 Z &= a_{21}R + a_{22}G + a_{23}B \\
 &+ a_{24}RG + a_{25}RB + a_{26}GB \\
 &+ a_{27}R^2 + a_{28}G^2 + a_{29}B^2 + a_{30}
 \end{aligned} \tag{1}$$

$$\begin{aligned}
 X &= a_1R + a_2G + a_3B \\
 Y &= a_4R + a_5G + a_6B \\
 Z &= a_7R + a_8G + a_9B
 \end{aligned} \tag{2}$$

$$\begin{aligned}
 X &= a_1R + a_2G + a_3B + a_4R^2 + a_5G^2 + a_6B^2 \\
 Y &= a_7R + a_8G + a_9B + a_{10}R^2 + a_{11}G^2 + a_{12}B^2 \\
 Z &= a_{13}R + a_{14}G + a_{15}B + a_{16}R^2 + a_{17}G^2 + a_{18}B^2
 \end{aligned} \tag{3}$$

We experimented with this formulation as well, as a basic linear transformation shown in Eq.(2) and a simplified version of the polynomial described in Eq.(1), where the constant and interaction terms for the tristimulus values were dropped as given in Eq.(3).

For each of the colour charts, the above formulations were tested using both least squares optimisation (‘pure’) and weighted least squares based on the content of the image being processed (‘heuristic’). Specifically, for each pixel the Euclidian distance from every patch was determined. Each patch’s weighting was based upon the total number of pixels that were closest to it in the camera’s device dependent RGB colour space. This process is analogous to preselecting colour patch values on the basis of the application area. Cheung and Westland [10] note that it would be sensible to select a characterization set based on the colour of the object being measured. Our approach is to augment this process by skewing the optimisation such that colours depicted in the image or more heavily favoured.

3.4 LEAF SEGMENTATION

Plant foliage may be used to identify a particular plant’s species and assess its health. Traditionally, these tasks would be carried out manually by an expert which is both time consuming and vulnerable to subjective variation. Automatic plant disease analysis using computers with/without supervising is nowadays widely considered. It, however, requires an image of a complete, isolated leaf as input to the process before further analysing other information in the leaf.

Extracting the leaf from outdoor images is difficult as the natural background contains undesirable things such as soil and other leaves. Also lighting, reflection and shape cause segmentation problems. We therefore propose a watershed framework using leaf characteristics as illustrated in Fig.5. First the background and leaf markers are created using colours, intensity and texture. The non-green areas detected using Otsu threshold [11] and the high detail areas, such as grass, measured using a local entropy are set as the background marker. The foreground markers are found using “opening-by-reconstruction” and “closing-by-reconstruction” to clean up the image and also connect blobs of pixels inside each of the foreground objects respectively [12]. Subsequently the markers are used as local minima in the gradient magnitude map of the greyscale image for the watershed transform.

The initial leaf extraction process gives good results for most images where the boundary of the leaf shows distinct gradients; however, in some cases the target leaves exhibit an unclear boundary resulting in protrusions, and in some cases the prominent veins inside the leaf cause partially extracted leaves. Therefore we introduce using characteristics of leaves to refine the results. Primary-secondary vein detection is applied with Hough transform [13] and protrusion-notch removal is processed via the graph of leaf contour where the perfect contour is assumed to fit with a polynomial of degree of 3. The unexpected plot is then replaced with the estimated fit curve. Consequently the protrusions are removed and the notches are filled.

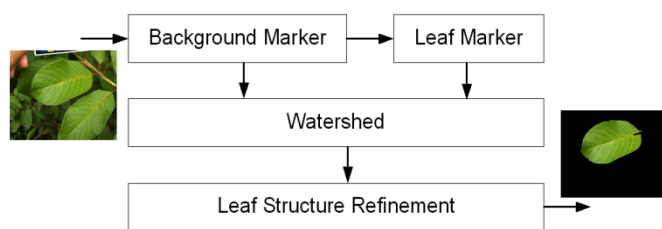


Fig.5. Leaf extraction process

3.5 DISEASE RECOGNITION

Once the leaf has been extracted, the margin, shape and venation can be estimated in order to identify the plant, whilst the colours of the segmented leaves can be utilized to determine the plant nutrition and health, also the visual symptoms of plant diseases. Machine learning technique can be employed here to develop a plant disease classification. In the early stage, the human experts are involved to a supervised learning as they can specify/label the diseases; as the result, the training data is set. Subsequently a model is formed. However, there are many diseases and also a new disease could occur. The process will be finished with the semi-supervised learning technique where the combination of the labelled data and new coming data is used.

4. RESULTS

The agricultural call center has been set-up and is operational at IITM's RTBI with the service support provided by our partners namely Tamil Nadu Agriculture University (TNAU), National Agro Foundation, Erode Precision Farm Producers Company Ltd., Dharmapuri Precision Farmers Agro Services Ltd. and Uniphore Software Systems. The services are initially being offered for 9 crops (Paddy, Groundnut, Brinjal, Turmeric, Tapioca, Mango, Sugarcane, Coconut and Tomato) to farmers in 3 districts Kancheepuram, Erode, Dharmapuri in Tamil Nadu. The farm data has been captured for 700 farmers. The preliminary results of the advisory services and disease recognition are presented below:

4.1 ADVISORY SERVICES

The first step to using the service was to create awareness amongst the farmers that such a service exists. In Kancheepuram district, 428 farmers were trained in between January to March, 2011 to use the service. 214 farmers called in to avail the advisory services during the period. Fig.6 presents the data on the type of queries received at the call center.

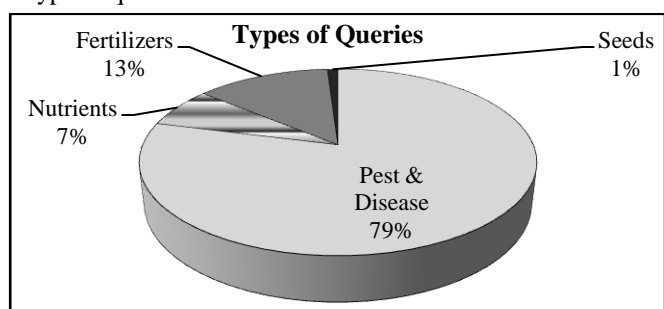


Fig.6. Preliminary Results of the Advisory Services

While some queries were related to information on seeds, fertilizers and nutrients, most of the queries were related to pest and disease mitigation. On analyzing these calls, it was found that a particular variety of paddy – ADT 37 – was prone to disease especially during the monsoon. 65% of the pest/disease related questions among paddy growers were pertaining to variety ADT 37, 19% to variety BPT, 12 % to Ponni and 4% to ADT 36. Details of all farmers growing this particular variety of paddy were retrieved from the database and mitigation steps were undertaken by sensitizing the farmers, extension workers and partners. Similarly, all farmers who were looking to cultivate paddy were made aware of the problems related to ADT 37 variety and were advised to sow a different variety – ADT 40/ADT 45.

4.2 COLOUR CORRECTION

While the agricultural call center started providing services to the farmers, work was taken up towards automatic disease detection. The data set presented here was taken from the University of Bristol's botanical gardens in late spring, 2010. Fifteen different plants were photographed with 5 different cameras including the DSLR (Canon EOS 5D mkII) which acts as our ground truth.

Example results for the botanical gardens are provided in Fig.7. Source and destination images are contrasted against the simplified quadratic transformation with and without the use of weighted least squares. Note that the transformed images look more similar to the target than the source image.

To obtain an objective measure of similarity we follow Rubner et al.'s approach [13] to assessing the distance between the transformed images and the target. Note that as the images are taken from different viewpoints it is not possible to perform a pixel-wise comparison. Rubner et al. demonstrate that their Earth Mover's Distance (EMD) metric used in conjunction with images transformed into CIE-LAB colour space effectively clusters perceptually similar images and hence is employed here. Global CIE-LAB colour histograms for the target, source, and transformed images were generated using 10 bins per channel. The EMD was calculated for the latter three images with respect to the target image. Note that as the target image is taken from a slightly different perspective to the others, one would not expect even 'perceptually perfect' correspondences to yield a zero EMD.

Fig.8 summarises the EMD results for out transformations compared against the EMD for the source image. In all cases the images are significantly 'closer' to the target after being transformed. For the botanical gardens' images the simplified quadratic formulation outperforms the linear transform for the Fujifilm Finepix F30 and the Apple iPhone 3G, whereas the linear transform is better for other two cameras. However, both produce faithful results and these numbers are in good agreement with our experience of eyeballing the images.

Encouragingly, our approach using weighted least squares optimisation generally performs slightly better than using the standard least squares approach.

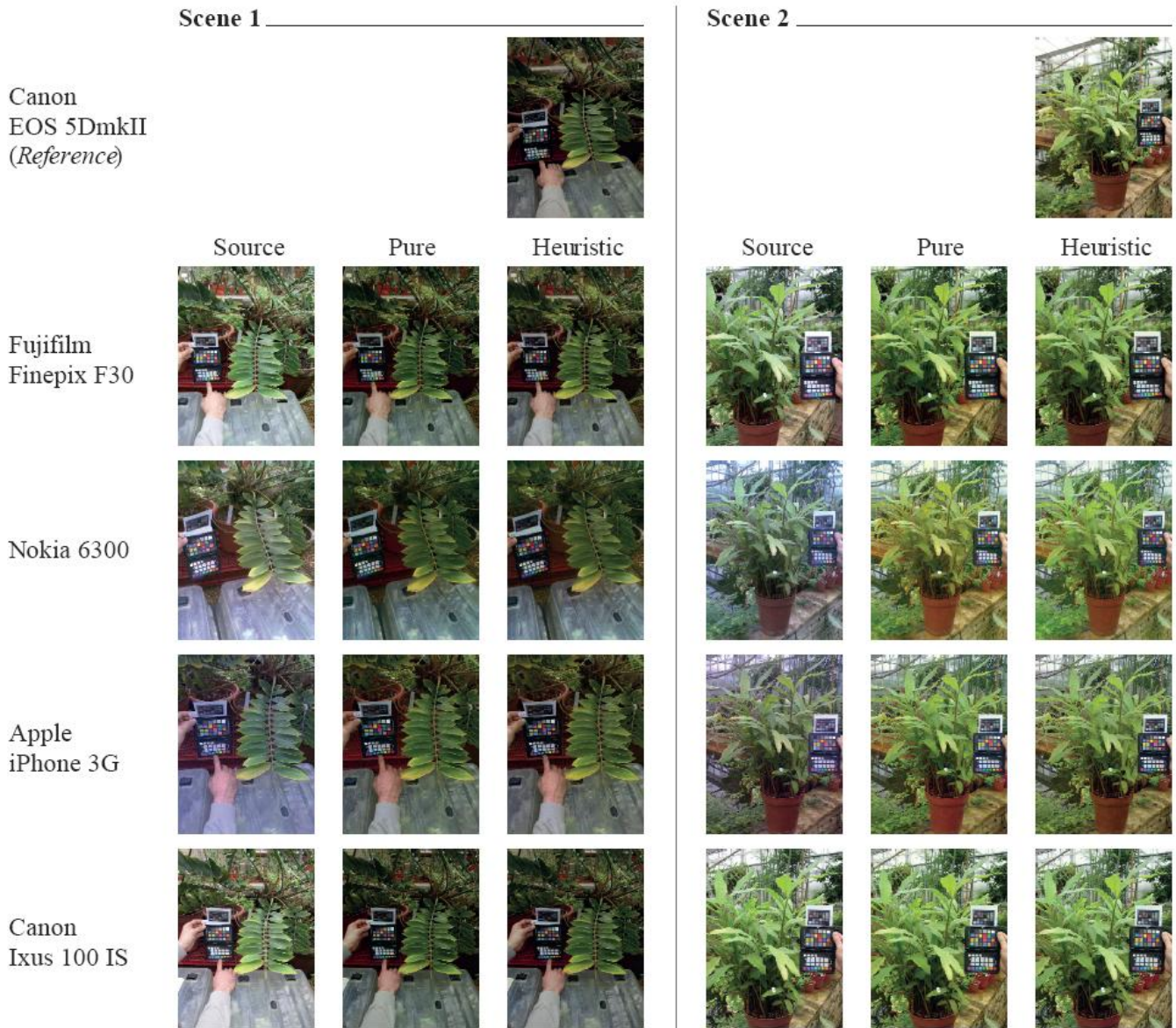


Fig.7. Sample scenes visually demonstrating relative efficacy of our methodology

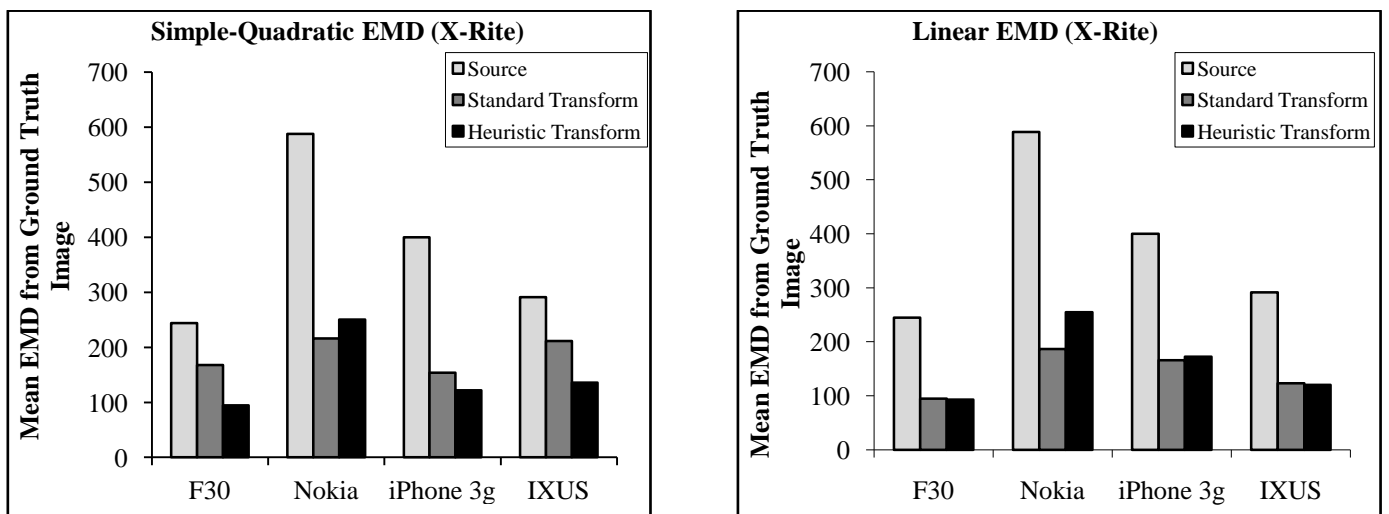


Fig.8. Mean EMD for all image analysed

4.3 LEAF SEGMENTATION

The proposed algorithm shown in Fig.5 was tested with 100 leaf images. The average precision and recall of our scheme are



Fig.9. Above: original images. Below: isolated leaves

5. CONCLUSIONS

The limited field trials show the importance of disease mitigation as a service for farmers. While there are direct and immediate benefits for farmers, this is an important step towards reducing crop losses in India. The work needs to be expanded and tested for extended period of time to understand the intricacies of the disease mitigation practices and fine tune it for maximum benefit.

At the same time, the early results make us more determined to work on a scalable solution. While the call center would be able to carry out disease mitigation, costs of extension workers would be a matter of concern. Automatic disease recognition will go a long way in reducing costs and provide the ability to scale. The information provided by farmers is used with the computer vision algorithms to assist the expert disseminating advice. An approach to standardising plant images for pathology diagnosis has been detailed. The efficacy of our custom colour chart and the use of a weighted least squares formulation have been demonstrated. It is anticipated that the method described here will be applicable to other applications utilising colour images that cannot be captured in repeatable way, for quality control and assessment purposes. Future work will more rigorously test our approach.

REFERENCES

- [1] http://en.wikipedia.org/wiki/Green_Revolution_in_India
- [2] Harsha de Silva, Dimuthu Ratnadiwakara and Shamistra Soysa, "Mobile phones to significantly reduce information-related transaction costs for smallholder farmers".
- [3] Claire J. Glendenning, Suresh Babu, Kwadwo Asenso-Okyere, "Review of Agricultural Extension in India - Are Farmers' Information Needs Being Met?" IFPRI Discussion Paper 01048, 2010.
- [4] <http://www.thehindubusinessline.in/bline/2008/05/01/stories/2008050150581300.htm>
- [5] Real Bergevin, Afshan Kinder, Winston Siegel, Bruce Simpson, "Call Center for dummies", 2nd edition".
- [6] Fairchild, M. D., "Color Appearance Models". Wiley-IS&T Series in Imaging Science and Technology, 2nd edition, 2005.
- [7] Luo, M. and Rhodes, P., "A study of digital camera colorimetric characterisation based on polynomial modelling," *Color Research and Application*, Vol. 26, No.1, pp. 76–84, 2001.
- [8] Tominaga, S., "Color coordinate conversion via neural networks," *Colour Imaging: Vision and Technology*, ed. Lindsay, 1999.
- [9] Cheung, T. L. V. and Westland, S., "Color camera characterisation using artificial neural networks," *Proceedings of the 10th Color Imaging Conference*, pp.117–120, 2002.
- [10] Cheung, T. L. V. and Westland, S., "Color selections for characterization charts," *Proceedings of the Second European Conference on Colour in Graphics, Imaging and Vision*, pp. 116–119, 2004.
- [11] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 91, No.1, pp. 62-66, 1979.
- [12] <http://www.mathworks.com/products/demos/image/watershed/ipexwatershed.html>
- [13] R.C. Gonzalez, R. E. Woods and S. L. Eddins, "Digital Image Processing Using MATLAB," Prentice-Hall, 2003.
- [14] Rubner, Y., Tomasi, C., and Guibas, L. J., "A metric for distributions with applications to image databases," *In ICCV*, pp. 59, 1998.