Decision Support Systems and Plant Phenomics

Valvachev, A. N. [1], Krasnoproshin, V. V. [2], Mikhalchenko, A.A, [3]
1) van_955@mail.ru, 2) krasnoproshin@bsu.by, 3) strrife@gmail.com,
Belarussian State University, 220030

Abstract: *The Earth is being polluted and territories that were fertile less than a decade ago now become barren. Meanwhile the human population is rapidly growing and the consumption of food is increasing exponentially. To overcome these challenges scientists are working hard on identifying quantitative phenotypes improving which would allow to overcome the consumption crysis. This paper gives a general overview and describes some practical problems in the field of plant phenomic and presents the algorithm to solve the problems described. In this paper we will consider only non-destructive automated phenotypung since exactly this field is the most interesting for us from the point of view of image processing and pattern recognition.*

Keywords**:** Image processing, phenomics, pattern recognition, plant phenotype.

1. INTRODUCTION TO THE PROBLEM

Since plant phenomics is a relatively new field of bioinformatics, we find it necessary to give an introduction not only to the problems we're solving, but to the entire field too.

That being said, plant phenomics is a field of bioinformatics that refers to a quantitative description of the plant's morphological and physiological properties. which are called phenoms.

There are two types on phenotyping, destructive (destructive phenotyping is an invasive measurement after which the plants can no longer exists and become unsuitable for further experiments) and non-destructive (measurements that do not harm the plant).

For the past years phenotyping has been progressing from the manual, whether non-destructive or destructive study to automated non-destructive study.

Not only in phenomics, but in ecology and botanics in general, one of the goals of the research is to observe the phenotypes of the organisms in various ecosystems, whether it would be a natural ecosystem or laboratory. Observation is held by taking measurements, usually of their appearance with time-lapse cameras, observing interactions of the organisms (in our field it would be the growth of plants). The data measured needs to be stored in a database for further analysis which usually boils down to comparing the measurements within a given period of time.

Research conducted in this field gave significant results and revolutionized the field of plant biology and agriculture because it allowed to make a high-throughput automated quantitative analysis of plant growth, development, productivity and stress resistance under variable conditions. For instance, analyzing the growth speed within different sets of plants given different fertilizers, we can make numerous conclusions about their effectiveness and the way they affect the growth of the plant.

Phenotyping research platforms has been extensively being established all over the world since 1990's. The scientific progress and the development of tech industry gave access to relatively inexpensive cutting-edge technologies, such as high-resolution cameras and high-performance hardware. These things are vital for plant phenomics, since it involves both laboratory- and field-based methodologies, using a number of non-invasive techniques, such as RGB visible imaging, imaging spectroscopy (multispectral and hyperspectral remote sensing), thermal infrared, fluorescence, 3D and tomographic imaging (MRT, PET and CT).

Indeed, a phenomical research requires a large set of measurements of the states of plants. In a good phenomical facility the amount of items of different species being researched is really large and can reach millions of units.

For these obvious reasons scalability and accuracy have always been a bottleneck of the entire field of science. Even experts in the field admin that it is really difficult to detect some important changes during the experiment for a human eye, whilst an automated computer-vision solution, though it may lose in some parameters, it would always win in accuracy and scalability.

The state-of-art phenomical solutions provide an entire phenomical complex with conveyor and numerous cameras taking pictures of plants from different angles, and the entire system is is driven by advanced enterprise software and the latest and most performant computer vision algorithms. Needless to say that such a complex and powerful solution is expensive and not affordable for most botanical institutions, which together comprise the force that drives this field of science.

Indeed, for the past decades, there's been a significant breakthrough not only in technical part of the problem but in botanical too. Various plant organs have been successfully phenotyped including shoot and root systems in hundreds of different species.

Thus, from the point of view of both industry and fundamental science, the development of phenomic as a field of science and applied phenomics such as phenomic analysis for juvenile arboreal plants will be really interesting.

2. PHENOMIC COMPLEXES AND IMAGE PROCESSING

There are a few commercial solutions to partially solve the problem of plant phenomics. For instance, the ScanAlyzer platform by LemnaTec is a plant phenotyping system that primarily focuses on the plant in its mature stage. The entire system is a greenhouse with conveyers on which the plants are automatically moved in front of stereoscopic cameras. The software provided by ScanAlyzer is proprietary and it analyzes the images to extract phenotypic-related information.

ScanAlyzer platform by LemnaTec is one of the most widely-spread and popular solutions on the market, fully developed and tested, though the platform is proprietary and expensive, and, therefore, requires a large investment in the infrastructure, and can no be easy deployed and maintained.

One of the main competitors of a ScanAlyzer is a custom growth phenotyping system developed by Optimalog. The main difference is the fact that Optimalog's system uses a robotic arm to position an array of sensors on top of a small plant within a growth chamber.

This is also a proprietary solution, and, just like in case of ScanAlyzer, there's only a limited amount of information about pricing and deployment. Nevertheless, the cost of the Optimalog's system is also high and not affordable for most of the Eastern Europe scientific institutions, since it requires a particular growth chamber design to accommodate the robotic arm. And, since it's a proprietary solution, maintenance expenses could comprise a significant amount too.

Alongside with these two phenomical solutions there's a many others, smaller and cheaper ones which, though, are usually more focused on some specific type of tasks and are not so heavily tested and qualitatively supported as their competitors listed above.

Belarussian botanical institutions investigate phenomc problems and make scientific breakthroughs, developing and researching plant phenomics, but they need something more suitable for the budget that would work and solve the problems that they are facing at the moment.

The purpose of this paper is to give a description and propose an algorithm and a solution based on state-of-art Computer Vision Open Source techniques that would give the simplicity of the acquisition and image processing system, providing results on some vital problems our botanical institutions are facing now.

Image processing algorithms and techniques with light sources from visible to near infrared spectrum provide non-destructive plant phenotype image datasets, and even the simplest digital cameras give good enough RGB images to process and analyze.

These approaches are commonly used in the field and have accelerated the quality and speed of real-time, high-precision phenotype data for modeling and prediction of plant growth and structural development. The application of combined image-processing technologies in phenomics and dedicated dynamically-controlled environments results in increasing performance and better results in the field of phenomics.

Image processing in plant phenotyping is not a exclusive and new research for detecting and recording qualitative and quantitative traits of objects (or, in our case, plants). Most of the studies in medicine, forensic, law-enforcement-related institutions and miscellaneous manufactures that deal with image processing are all about detecting and precisely calculating traits of objects, whether these would be detecting dimensional size or color or any other feature.

Since image processing methodologies are the key technologies in plant phenomics which, properly speaking, are rapidly developing, the main goal is to measure quantitative phenotype through the interaction between plants and light, such as reflected photons, absorbed photons, or transmitted photons, which actually means that we will process and analyze the images of the plants taken from different cameras in the greenhouse.

3. DECISION SUPPORT SYSTEM

A decision support system is a computer-based information system that supports business or organizational decision-making activities.

A decision support system usually help to partially automate and speed up the workflow by proposing a few most suitable options to the operator who is supposed to choose the best of them.

At the moment, plant biotechnologists and horticulturists do not have tools based on the image analysis to distinguish rooted and dead green cuttings during rooting in the soil substrate.

Removing dead green cuttings is vital part of optimization since freeing the space occupied by plants that have no chance to prosper again. Indeed, the soil they take and the fertilizers that are applied to them could be used much more efficiently on the new, prospering green cuttings.

Moreover, plants propagated in vitro by microclonning pass through vulnerable stage of adaptation and rooting in ex vitro conditions, when many of them fade and die.

4.PHENOTYPING PLATFORM AND THE ALGORITHM

The algorithm proposed in the paper is going to be used in the experimental phenomic complex established in a greenhouse in one of the Belarussian State University botanical facilities.

The environmental parameters there will be constant during rooting phenotyping. Success of rooting will be monitored manually at days 50 and 100.

The trays will be monitored by visible (RGB) imaging every 24 hours and to this purpose five RGB cameras will be installed used simultaneously to take pictures of the plants during the process of rooting. The cameras will be placed under different angles and distances to cover all green cuttings in each tray.

Fig. 1 — Example of the input image

There will be 5 cameras that would capture the plant from al required positions: front, right, left, back and top. Original images from each camera will be taken and stored in databases and then analyzed using pattern recognition and image analysis algorithms

Here, we propose to develop a platform for phenotyping stem cuttings and juvenile clones used for vegetative propagation of ornamental woody plants.

The degree of fading can be detected with the human eye and is mostly about the color of the green cuttings. For instance, if the plant hadn't received any water during the period, its leaves' color would move from the green color range closer to the yellow or even brown color range.

For instance, on the Fig. 1 we can observe a set of leaves of grass one of which that is thin, bending to the side and definitely does not look healthy, as detected analyzing the RGB-spectrum, emits more shades of brown that other leaves.

The described problem is one of the easiest amongst that we solve. Indeed, after separating each of the cuttings we should only take measurements of the percentage of some particular color in the RGB spectrum of the image.

The more challenging and interesting problems are the actual splitting of the leaves (i.e. detecting the sample t which this or that leave belongs to) and detecting the growth dynamics.

The solution to both of these problem is based on detecting the contours of each leaf (or root, as would be show below) and find the median.

But, obviously, the very first step would be cleaning the image, i.e. removing the background and applying the dimensional transform to level the angle under which the picture was taken.

Each container is going to be marked with a QR-code, just like it's done in the most advanced phenomic systems like ScanAlyzer.

On the image below you can see the container in which there are a few cuttings with bare roots. This particular image belongs to a the set where the main goal was to detect the length and growth speed of the root system of the plant.

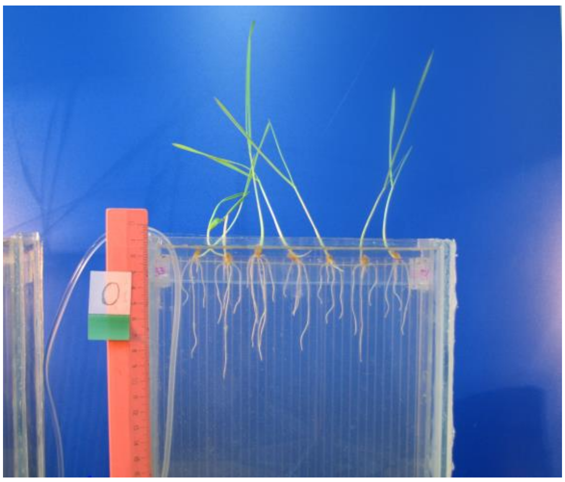

Fig. 2 — Cuttings with bare roots

So, the entire process comprises of the following steps after retrieving the original image and reading the QR code from the container to be able to get the information about the plan:

1. Removing the background. The images of plants will be separated from the background and stored as individual images in the dedicated database. Since tissues do not pass and do not reflect the blue light, the blue spandex fabric will be used as a background when taking pictures;

2. Based on he position of markers placed on the corners of the container the dimensional transformation which was applied due to the positioning of the camera would be detected and apply the inverse transform to the image would be applied. On Fig. 3 you can see the result of transformation of the binary mask retrieved on step 1.;

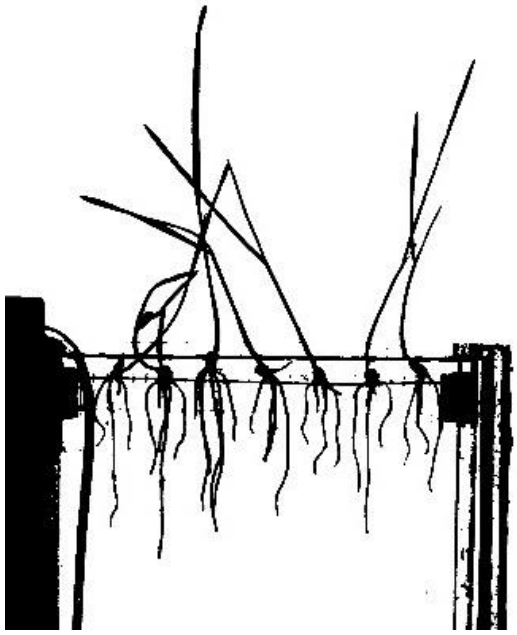

Fig. 3 — Binary mask of the transformed image

3. Images of stems (non-photosynthesizing organs) would be separated from leaves (photosynthesizing organs) based on difference in their spectral properties, since photosynthesizing organs absorb hundred times less in green/yellow part (500-600 nm) of the visible electromagnetic spectrum than in blue and red parts (400-500 nm and 600-700 nm).

Fig. 4 — Green spectrum separation

The example of such separation of leaves can be seen on Fig. 4.. These stem-free images would also be stored in the dedicated database.

4. based on the color difference the leading lines of the plants will be detected and individual plants would be separated one from another for further analysis. Background-free and stem-free images would also undergo this procedure;
5. For each sample, detect the percentage of the shades of brown in the leave's RGB spectrum. These data will be presented as spectral bar histograms of distribution of intensities which would be stored in the histogram database and monitored on the daily basis to detect the dying of plants. Patterns of averaged specific histogram changes related to death, survival (callusogenesis) or rooting will be identified and stored. They will be used as patterns for early recognition of physiological fate of green cuttings and making recommendations to horticulturists. Obtained patterns will be used for monitoring and predicting rooting effectiveness during three sequential seasons;
6. For each individual plant its size would be detected (i.e. dimensional measurements). This can be done because the container size is well-known (or can be loaded with the help of container ID) and the dimensional transformation is applied;
7. Based on the operators decisions the system would adapt and correct its behavior to identify individual plants and its characteristics better in future.

5. CONCLUSION

A brief description of plant phenomics and main problems and difficulties in the field was given in the paper, some particular problems were highlighted and ways to solve them were proposed.

Research in plant phenomics is beneficial for both enterprises and non-profit companies and the research itself continues to benefit not only from the researches held in botanics and other related fields but first of all from development of cutting-edge image processing technologies and breakthroughs in the engineering.

This advanced research enables observation of important phenotypic traits and how these traits change depending on environment and genotype.

The amounts of data stored would grow linearly with a hight speed, and therefore, it is necessary to manage and process data efficiently., and again, high results achieved in Computer Science (in particular, BigData) give a significant impulse to the solution of problems in phenomics.

There is an urgent need to develop more adaptable, less expensive and sophisticated data analysis infrastructures for the problem of phenomics for analyzing high-dimensional phenotype datasets. These achievements will support the speedup and improvement of plant growth, development, or responses to adverse environments, which is so vital due to the consumption crysis which we're about to face in relatively near future according to some experts.

The study of this problem is not complete and will be continued in future in order to complete the main objective is to develop and introduce a fully operable phenomic complex that would be used in Belarussian botanical institutions.

6. REFERENCES

1. Aerts H. J., Velazquez E. R., Leijenaar R. T., Parmar C., Grossmann P., Cavalho S., et al. (2014). *Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach.* Nat. Commun. 5 4006 10.1038/ncomms5006
2. Brown TB, Cheng R, Sirault XRR, et al. *TraitCapture: genomic and environment modelling of plant phenomic data.* Curr Opin Plant Biol. 2014;18:73–9.
3. Eberius, M.; Lima-Guerra, J. (2009), "*High-Throughput Plant Phenotyping – Data Acquisition, Transformation, and Analysis*", Bioinformatics 7 (1): 259–278, doi:10.1007/978-0-387-92738-1_13, ISBN 978-0-387-92737-4 pp.123-127.
4. Houle, David; Govindaraju, Diddahally R.; Omholt, Stig (2010). *"Phenomics: the next challenge". Nature Reviews Genetics 11* (12): 855–866. doi:10.1038/nrg2897. ISSN 1471-0056. press/accepted.
5. Brown, Lester R. (2011). *World on the Edge: How to Prevent Environmental and Economic Collapse.* Earth Policy Institute. Norton. p. 7. ISBN 113654075X.
6. Marius Cioca, Florin Filip (2015). *Decision Support Systems* - A Bibliography 1947-2007.