

Analysis of Pattern Formation on the Surface of Respiring Leaves¹

Aaron Luttmann, Emily Stone and Johnathan Bardsley

*Department of Mathematical Sciences
University of Montana
Missoula, MT 59812*

Abstract

The process of photosynthesis is facilitated in a plant by pores in its surface called *stomata*. When a particular stoma is open, CO₂ is absorbed through its aperture, but H₂O is also lost due to evaporation. Thus a plant will seek a stomatal aperture that balances its need for CO₂ with its aversion to H₂O loss. In order to visualize a particular leaf's stomatal aperture distribution and how it changes with time, fluorescence data is collected at regular intervals as digital images, resulting in a video sequence. In observing such data, plant biologist have noticed - in over 200 plant species - that, even under spatially varying environmental conditions, stomatal apertures are often synchronized into spatially extended patches. This behavior is observed even though patchiness often results in less than optimal local CO₂ uptake for fixed water loss. It is not understood why stomatal patchiness occurs, and it is the objective of this paper to shed light on this phenomenon via the analysis of one such fluorescence video. In order to perform our analysis, we first extract the patches from the video via a 3D, PDE-based segmentation method. With the segmented video in hand we then analyze the patches and their dynamics using a variety of measures.

Key words: Image Segmentation, Archetypes, Mutual Information

Email addresses: luttman@blc.edu (Aaron Luttmann), stone@mso.umt.edu (Emily Stone), bardsleyj@mso.umt.edu (Johnathan Bardsley).

¹ This work was supported by the NSF under grant DMS-0504325 and by Montana NSF EPSCoR.

1 Introduction

In this paper we analyze video data from an experiment performed at the *Complexity and Stomatal Behavior* research lab at Utah State University [23]. The lab studies the spatial dynamics of leaf respiration, where a complicated evolution of fluorescing patches on the surface of leaves indicates a non-trivial process of communication and synchronization of the central organ of plant respiration, the stomatal pore. The dynamics of this process was initially analyzed in Peak et al. [23], and here we present analysis that preserves the spatial structure of the domain. Both the development of the techniques for this analysis, and the result applied to the experiment, are the topics for discussion in what follows.

Plants require CO_2 to engage in photosynthesis, the process by which they convert solar energy to food. A plant absorbs CO_2 through pores on its surfaces called *stomata*. When the stomata are opened, however, H_2O is lost through evaporation, which is detrimental to plant function. Thus stomatal apertures must be regulated in such a way that the plant's need for both CO_2 and H_2O is balanced. Peak, et. al. [23] make the following statement regarding this process: "A central paradigm of plant biology is that, in the face of spatially heterogeneous and temporally varying environmental conditions, a plant continually adjusts its stomatal aperture so that, over time, it maximizes CO_2 uptake for a fixed amount of water loss." It is thought that this occurs over the entire plant rather than for each stoma individually and that the stomata adjust their apertures based on external environmental conditions. However, the fact that stomatal apertures are often synchronized into spatially extended patches, a phenomenon observed in over 200 plant species [1], suggests that neighboring stoma communicate. This claim is further supported by the fact that stomatal patchiness has been observed even when it results in less than optimal local CO_2 uptake for fixed water loss [2].

The data analyzed in this paper is from a visualization of the respiration process of a cocklebur plant (*Xanthium strumarium L.*). In order to visualize the stomata opening and closing, a chlorophyll dye is injected into the leaf. This dye fluoresces when interacting with CO_2 , and an image of this fluorescence is collected by a charge-coupled device (CCD) camera. Each image consists of a 512×512 array of pixels. Regions of high relative intensity correspond to areas in which the stomata are closed, whereas inter-vein regions of low relative intensity correspond to areas in which the stomata are open. The images in the sequence were collected every 40 seconds over a period of several hours. (For more details on the data collection procedure and chlorophyll fluorescence, see [23,19,10].) A video of the data can be found at http://hans.blc.edu/~luttman/movies/original_half_set.avi. Four sample images from the video sequence are shown in Figure 1. In frames 150,

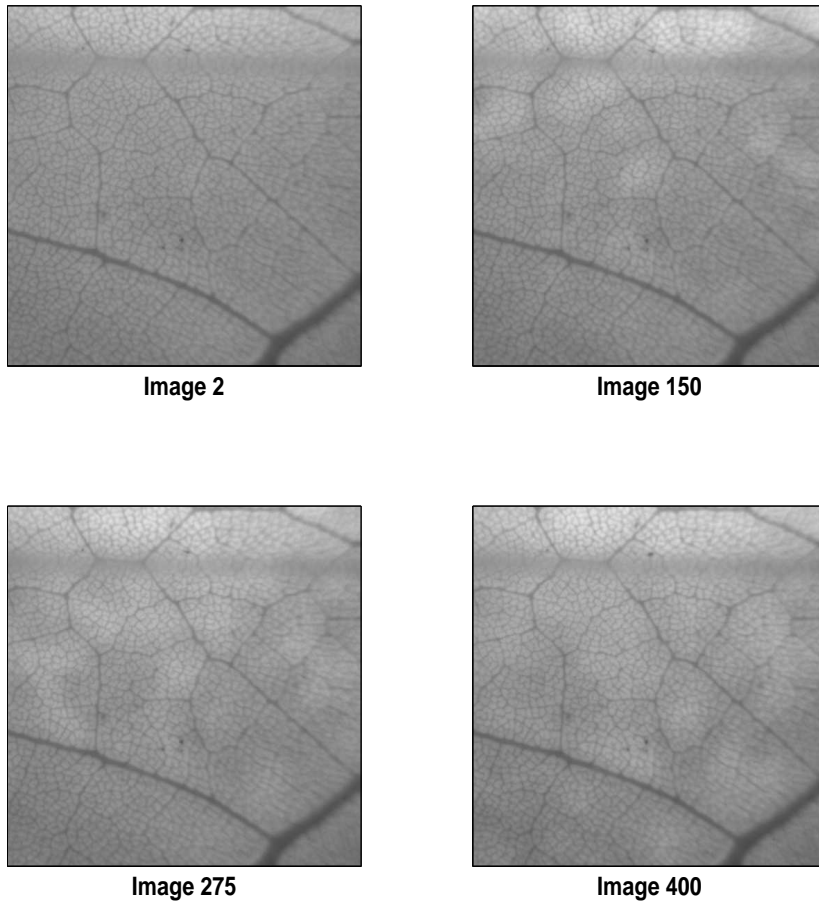


Fig. 1. Four 512×512 Images from the Video Sequence of 436 Frames.

275 and 400, the light (closed stomata) and dark (open stomata) patches are evident.

The video data must be processed in order to identify coherent patches, and the end result is a binary segmentation of the image into black and white regions. This is achieved with a PDE-based, variational method that operates over three dimensions – the two dimensional extent of the frame plus time (advancing frame number) – generating a segmentation that varies smoothly in time. A critical first step is the preprocessing of the images, where instrumental and experimental noise is removed. The methods presented here are part of the current state-of-the-art in image processing [17].

Once the fluorescence patches are extracted, their dynamics can be analyzed. After an initial visual inspection, the domain is separated into regions delineated by veins. Each contains one or two patches that evolve over time. The

position and size of the patches can be analyzed directly, but of real interest is the dynamics of the shapes of the interacting patches. These are studied by first a compression with PCA (Principal Component Analysis) and subsequent Archetypal Analysis of the patterns. This method separates the data into space and time components, and allows for the analysis of the development of patches in one region compared with that of another region. Time series methods were used in this analysis, including the construction of two-dimensional, pair-wise mutual information for all the regions. From this we are able to objectively determine the relationship of dynamics of patches in one part of the leaf to another. We found apparent coupling between neighboring regions, with some caveats, discussed in detail later in the paper.

The development of these techniques for dynamic two dimensional pattern analysis is one aim of this project. We demonstrate here a method which objectively collapses the complicated dynamics of the patches to a discrete time series and corresponding pattern bases.

The paper is organized as follows. In Section 2 we outline the experiment performed by Mott et al. [23], and the form of the video data. The video preprocessing and segmentation algorithms are presented in Section 3. Analysis of the resulting segmentation appears in Section 4, and in Section 5 we summarize our results and outline future work.

2 The Experiment

In order to better understand the way in which a leaf's stomatal apertures are regulated, it is natural to visualize the process. This can be done using a technique known as chlorophyll fluorescence, used for visualizing photosynthesis. This process is explained in detail in [10,26], but we outline the main ideas here.

A *Xanthium strumarium L.* (cocklebur) leaf is placed in a light-penetrating chamber such that temperature is controlled and gasses can be independently regulated and measured both on the top and bottom surfaces of the leaf. A dye is injected into the leaf, so that it fluoresces when its stomata close. The fluorescence happens at wavelengths above 700 nm, and is recorded using a camera with a long-pass filter that measures light in the appropriate spectrum. Stomatal dynamics are initiated by decreasing the concentration of H_2O on the upper surface of the leaf. This causes the leaf to close its stomata, since a decrease in humidity will correspond to an increase in H_2O loss through evaporation. The closing of the stomata causes a decrease in the amount of CO_2 absorbed, resulting in higher concentrations of CO_2 available for interactions with the fluorescent dye and therefore increased measured fluorescence.

Thus there is an inverse relationship between photosynthesis and fluorescence. Since the density of stomata ($\approx 20,000$ per cm^2) is of the same order of magnitude as the density of image pixels ($\approx 40,000$ per cm^2), it is reasonable to assume that the measured change in intensity at each pixel roughly corresponds to local changes in stomatal aperture. In this way, fluorescence can be interpreted as an inverse measure of average stomatal aperture size.

As the stomatal apertures of the leaf change size, images of the fluorescence can be measured (in the experiments analyzed here the images are taken approximately every 40 seconds). The measured data is then a gray scale video composed from these still images. Each experiment lasts approximately 6 hours, resulting in videos on the order of 800-1000 images. All experiments were performed by the *Complexity and Stomatal Behavior* research lab at Utah State University; further details can be found in [23]. Several example images from an experimental run can be seen in Figure 1.

The most commonly observed result of the experiment is that stomatal apertures across the entire leaf adopt a uniform size in direct reaction to environmental conditions, and this steady-state is usually achieved quite quickly (approximately 30-60 minutes). It is sometimes observed, however, that spatial groups of stomata will synchronize their apertures, forming “patches.” The stomata in these patches then act as a unit, opening and closing in synchrony. Moreover, each patch can change size and position in time. In most cases, after 1-2 hours of oscillations in patch size and shape, the uniform steady state mentioned above is reached, i.e. all stomata adjust to a uniform aperture size. On rare occasions, however, no steady state is found for as long as the data is measured (approximately 6 hours). This suggests that the solution state for the global optimization is unstable and that the leaf is sensitive to microscopic changes in gaseous configurations. The data set we analyze here is from such a regime.

3 Video Segmentation

Our first task is to provide a binary segmentation of the video into light and dark regions. For this, we follow the general approach outlined in [17], with modifications that yield a segmentation that is more appropriate for our analysis. The video must be preprocessed before the segmentation is performed. In the preprocessing step, the task is to remove both instrumental and experimental noise. A background model is then subtracted from each frame, and the video is rescaled in order to emphasize the contrast between the light and dark regions. In order to extract the fluorescing and non-fluorescing regions from each frame of the video, we use a variational PDE-based segmentation method [17]. The segmentation is performed in three-dimensions (over the

entire video sequence) so that the algorithm is able to globally compare the changes between the video sequence and the background model. In contrast, if the segmentation is done in two-dimensions, i.e. frame by frame, differences between the images and the background model tend to be overemphasized in frames with little or no dynamics. Such frames occur, for example, in the early stages of the video.

3.1 Preprocessing

First, each image in the video sequence is denoised using the total-variation regularized L^1 minimization procedure proposed by Chan and Esedoglu in [7]. This is a variation of the well-known ROF denoising scheme [24]. Next, each image is normalized against global lighting effects caused by variations in experimental lighting conditions, and is linearly scaled so that the intensity values in the sequence take up the full range $[0, 255]$. The resulting images, corresponding to the original images in Figure 1, are shown in Figure 2. Notice that the stomatal patches in Figure 2 are much clearer. For more details on the denoising and normalization algorithms see [17] and the references therein.

An additional preprocessing step is required for the segmentation of the video. Since the veins are not to be extracted and their intensities are approximately constant throughout the video, they can be removed by subtracting a static background model, $B(x, y)$, generated from the experiment. Let $I(\vec{x}) = I(x, y, z)$ denote the normalized video with pixel coordinates x and y and frame coordinate z . The background subtracted data is then defined

$$D(\vec{x}) = |I(\vec{x}) - B(x, y)|. \quad (1)$$

Large values of $D(\vec{x})$ correspond to the regions of the video that differ substantially from the background model; these are the regions of active fluorescence. In order to ensure that these regions have sufficiently smooth boundaries, $D(\vec{x})$ is convolved with the 2D Gaussian kernel with $\sigma = 3$. We will denote the resulting, smoothed data also by $D(\vec{x})$.

The final preprocessing step is to nonlinearly scale the smoothed data as follows:

$$D(\vec{x}) = \frac{1}{\pi} \arctan (s_1(D(\vec{x}) - s_2) + 0.5),$$

where s_1 is a scale factor for the enhancement, and s_2 is a lower bound on the intensity values to be brightened. For the data seen here, $s_1 = 15$ and $s_2 = 0.33$. This enhances the bright regions being extracted.

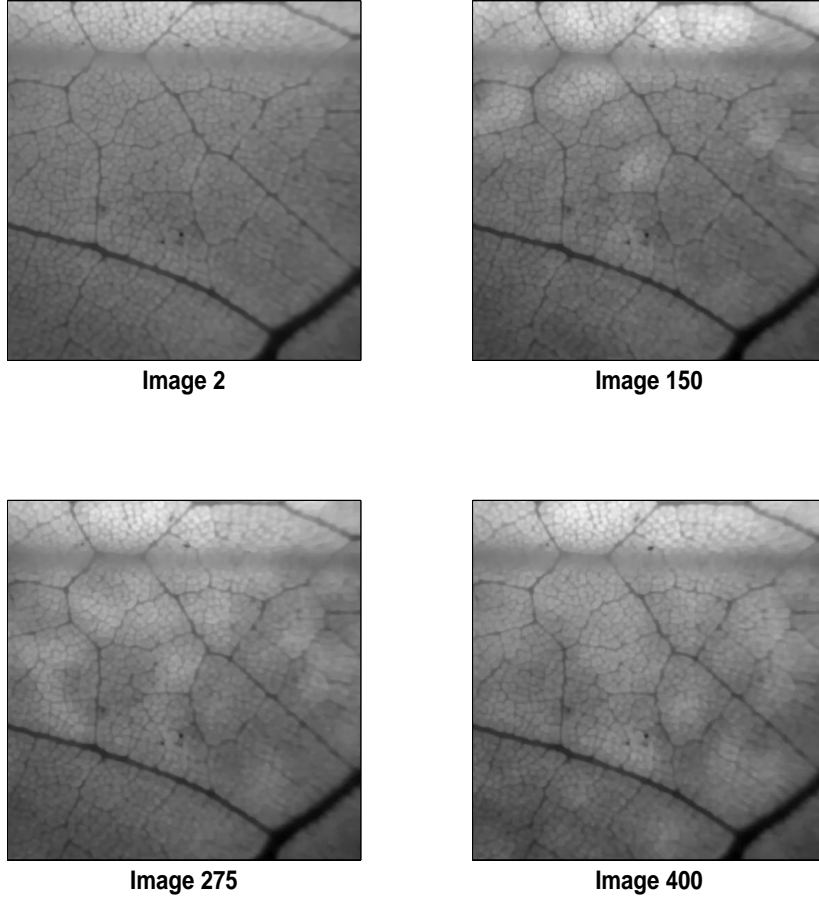


Fig. 2. Denoised and Normalized Images.

3.2 Segmentation

We now give a brief outline of the segmentation algorithm presented in [17], where the desired segmentation is obtain, following [13], by minimizing

$$E_{\text{MV}}(\Phi) = \beta \int_{\Omega} |D - c_1|^2 (1 - H(\Phi)) + |D - c_2|^2 H(\Phi) d\vec{x} + \alpha \int_{\partial\Omega} g(\vec{x}) dS. \quad (2)$$

Here Ω is the video domain; c_1 and c_2 are the mean values of D on $\Omega_1 = \{\vec{x} \in \Omega \mid \Phi(\vec{x}) < 0\}$ and $\Omega_2 = \{\vec{x} \in \Omega \mid \Phi(\vec{x}) > 0\}$ respectively; H is the Heaviside function; and α and β are weighting factors whose ratio is the relevant quantity. The first two terms of (2) are the fidelity terms that ensure that the result will approximate D in $L^2(\Omega)$, and the last term is a *geodesic active contour* term [6], which ensures that the boundaries of the segmentation are smooth.

Here g is an edge-indicator function computed from the denoised video given by

$$g(x, y, z) = \frac{1}{1 + I_x^2 + I_y^2}. \quad (3)$$

Note that, since edges are two-dimensional features, the edge-indicator function is computed on each frame independently. The problem of minimizing (2) can in turn be recast as an evolution equation of the form

$$\frac{\partial \Phi}{\partial t} = \alpha \operatorname{div}(g \nabla \Phi) + \beta(c_2 - c_1) \left(D - \frac{1}{2}(c_1 + c_2) \right), \quad \Phi(\vec{x}; 0) = \Phi_0(\vec{x}). \quad (4)$$

Thus, by numerically solving (4), we will obtain a segmentation of the video. For a more detailed derivation of (4), see [17] and the references therein.

We solve (4) numerically using an explicit discretization in time and a semi-implicit discretization in space. Using the natural pixel discretization of the video domain Ω , we have grid points (x_i, y_j, z_k) in the video sequence, where j runs over the columns in each image, i runs over the rows, and k runs over the frames. Also $\Delta x = \Delta y = \Delta z = 1$. We denote the function u at grid point (x_i, y_j, z_k) and time step n by $u_{i,j,k}^n$ and define

$$A_s = \frac{\partial}{\partial s} g \frac{\partial}{\partial s}$$

for $s \in \{x, y, z\}$. Then the semi-implicit discretization of (4) is given by

$$(\mathcal{I} - \alpha \Delta t (A_x + A_y + A_z)) \Phi^{n+1} = \Phi^n + \beta \Delta t (c_2^n - c_1^n) \left(D - \frac{1}{2}(c_1^n + c_2^n) \right), \quad (5)$$

where \mathcal{I} is the identity operator, and the operators A_x , A_y , and A_z are approximated using a finite-volume approach (cf. [18]). Note that g and D are independent of t , but that c_1 and c_2 depend on the level set function and therefore, implicitly, on t .

The inverse on the left in (5) is approximated using the *additive operator splitting* (AOS) method proposed in [29]. The AOS time update for (5) is given by

$$\Phi^{n+1} = \frac{1}{3} \sum_{s \in \{x, y, z\}} (\mathcal{I} - 3\alpha \Delta t A_s)^{-1} \left(\Phi_{i,j,k}^n + \beta \Delta t (c_2^n - c_1^n) \left(D - \frac{1}{2}(c_1^n + c_2^n) \right) \right). \quad (6)$$

By comparing the Taylor expansions for this and the semi-implicit scheme, it can be seen that the methods in (6) and (5) differ by a factor of $\mathcal{O}(\Delta t^2)$. Each of the operators $\mathcal{I} - 3\alpha \Delta t A_s$ is symmetric, positive-definite, and tridiagonal, so the efficient Thomas algorithm can be used to compute each of the linear system backsolves. Moreover, each of the backsolves is independent of the others, so they can be, and were in this case, computed in parallel. For further details of the numerical scheme see [17] and the reference therein.

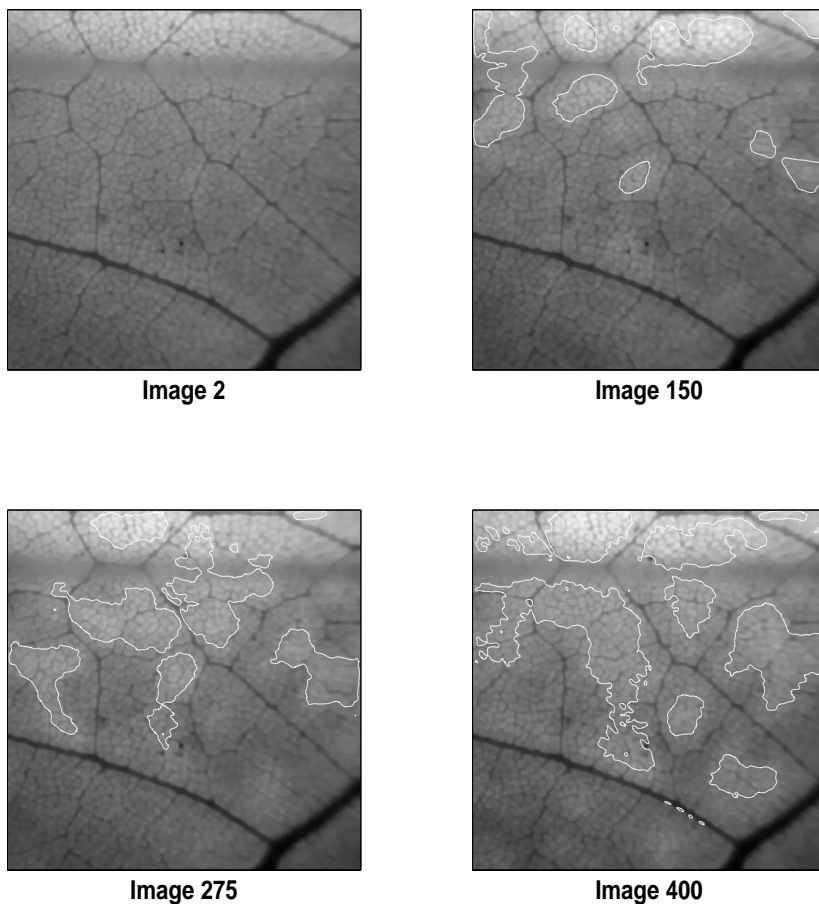


Fig. 3. Segmented Images.

We obtained the desired segmentation of the leaf video by computing the 15th iterate, Φ^{15} , of (6), with $\alpha = 1.0$ and $\beta = 2.0$. In Figure 3, the zero level set of Φ^{15} is overlaid on the preprocessed data. The white contours separate the regions of active fluorescence from the regions that are not fluorescing. Note that no contours are present in image 2, correctly reflecting the fact that there is no fluorescence occurring in that frame. This fact is not captured correctly when the segmentation is performed in 2D on each image individually. In Figure 4, the corresponding binary segmentation of the leaf video frame is shown. Black pixels correspond to grid points $\vec{x}_{i,j,k} = (x_i, y_j, z_k)$ such that $\Phi_{15}(\vec{x}_{i,j,k}) < 0$, and white pixels to points such that $\Phi_{15}(\vec{x}_{i,j,k}) > 0$.

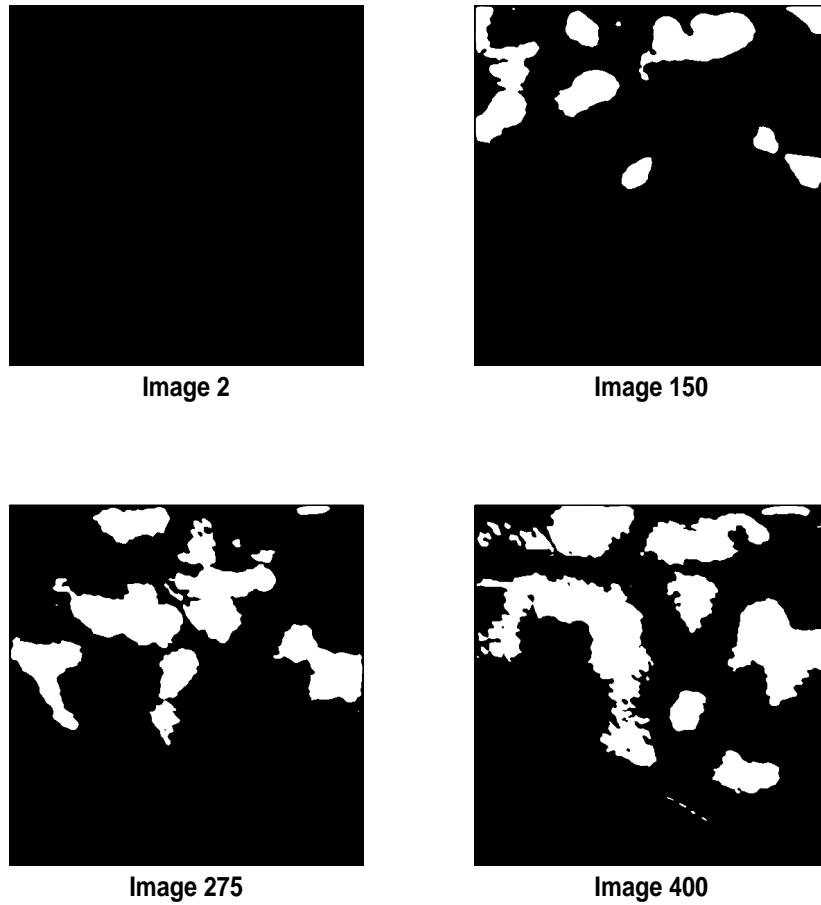


Fig. 4. Binary Images.

4 Patch Dynamics Analysis

We proceed with the analysis in order of increasing detail. We begin with dynamics measurements at the frame level. We then divide each frame into subregions, and analyze the dynamics at the subregion level. Finally, in order to explore the correlation between the dynamics in different subregions, we separate the patch dynamics into shape and time characteristics using a hybrid PCA/Archetypal analysis technique.

4.1 Frame Level Dynamics

The large scale information about the domain concerns the overall extent of the fluorescing patches: how many pixels are “on” in any given frame (see

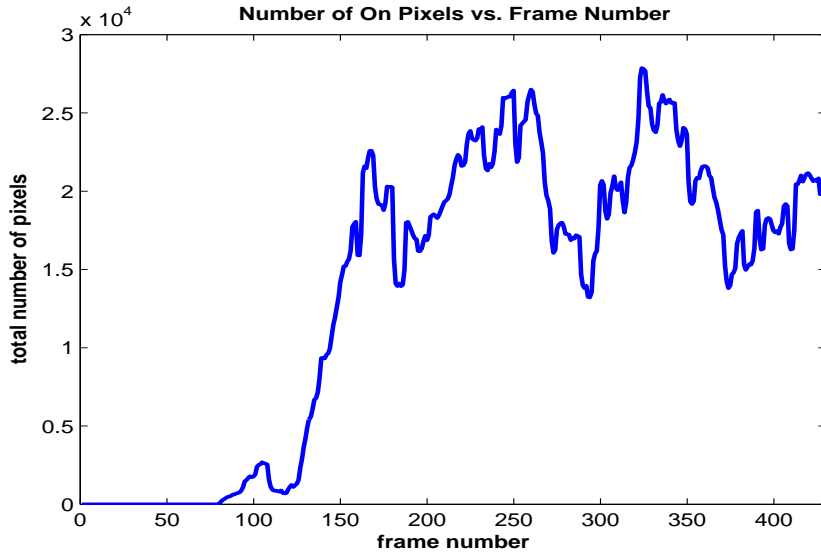


Fig. 5. Number of ON pixels in whole domain vs. frame number.

Figure 5). The visual intuition that the patches spread across the domain and then vary in complicated manner until the end of the data set is confirmed by this time series. The initial part of the data set is retained here, we see that until 50-100 frames the leaf is quiescent, no patches exist.

Next we consider the number of patches seen at any given time. This is computed via a simple numerical algorithm that counts pixels within a given patch, and logs the number of patches encountered over 100 pixels in size. The graph in figure 6 demonstrates an evolution similar to that seen in figure 5. This is as expected, since if there are more patches there are more “on” pixels. We should point out that these analyses would not have been possible using the original, i.e. non-segmented, data.

4.2 Subregion Level Dynamics

The fact that coherent patches form at all is indicative of local coupling between stomata, so our remaining analysis concerns the extent of long range coupling, or communication, between the coherent patches. This is difficult if not impossible to determine by visual study of the video; the patch dynamics are complex and occur in widespread areas over the leaf domain. After some inspection, however, it becomes apparent that patch movement is mainly restricted to regions that are bounded by veins of an intermediate size. We therefore analyze the dynamics in the regions separately, and then compare the dynamics within each region to all the others. The regions were delineated by hand and masks were formed to isolate the image in a single region from the rest of the leaf domain. A region map is plotted in figure 7. Note that the

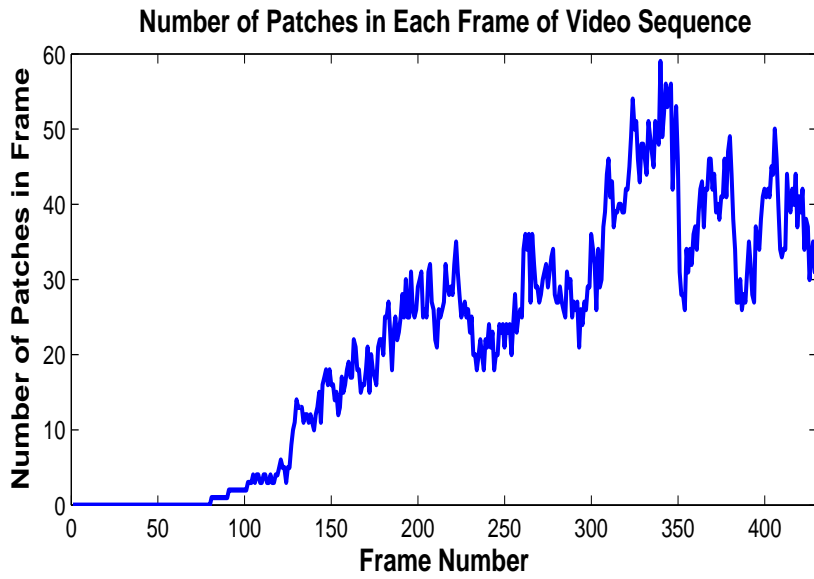


Fig. 6. Number of patches over 100 pixels in area vs. frame number.

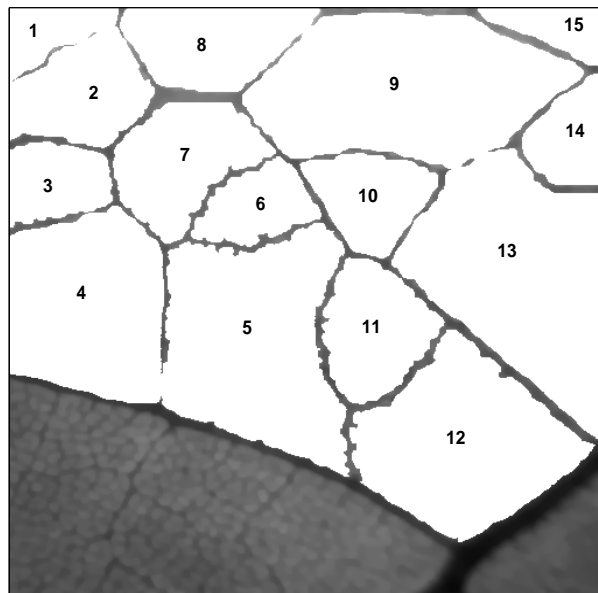


Fig. 7. Leaf Region Masks.

area at the bottom of the frame, where no regions are defined, is where no patches form in the experiment.

Within each region patches are seen to appear and meander around, changing position, size and shape. We analyze the position and size by computing patch size and center of mass for each region for each frame. This was done by counting the number of white pixels (to give size) and the center of mass of

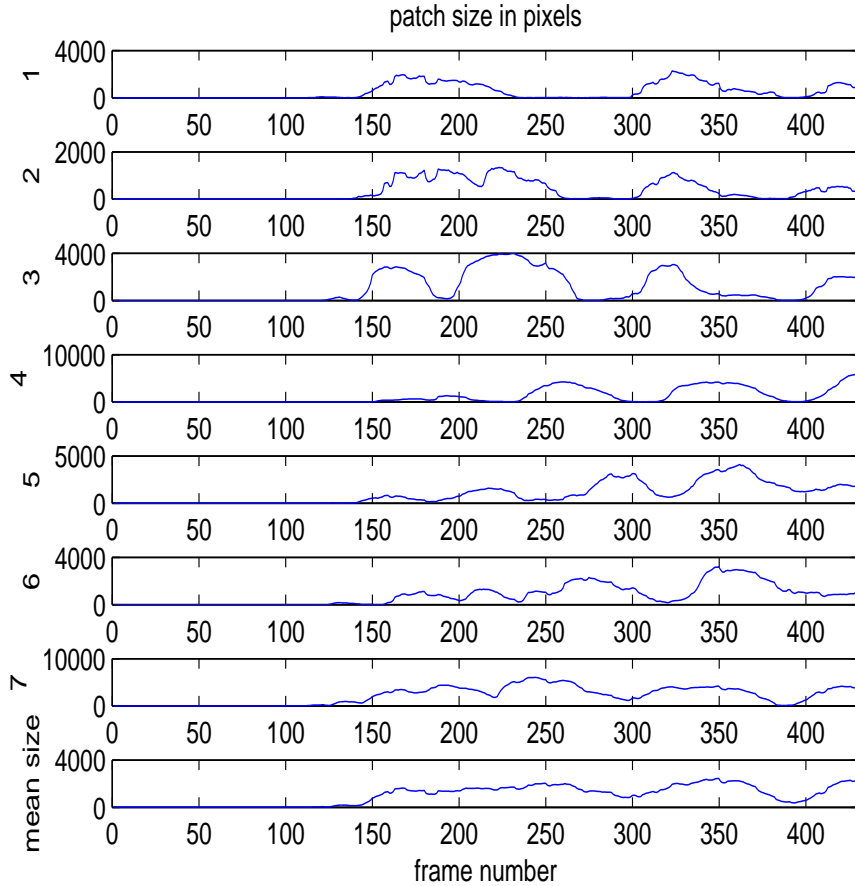


Fig. 8. Patch Size in number of pixels vs. frame number for the first seven regions. Last time series is the average patch size for all 15 regions vs. frame number.

the corresponding patch. A plot of the patch size vs. frame number for the first 7 regions is shown in figure 8. We see that for the first 120 or so frames the leaf stomata in this region are static. After this the patches form and change until the end of the data set. Beyond that the patches appear and disappear in more or less coordinated fashion, an issue we take up in the next portion of the analysis. The bottom time series is of the average patch size for the whole domain vs. time, which varies much less dramatically, as would be expected.

The possibility that the patch size for a given region is normally distributed (removing the first 150 frames from the data when the patches are initially forming) was tested with the Jarque-Bera test (using Matlab’s `jbttest`). This yielded p-values on the order of 10^{-5} for all regions, allowing us to reject the hypothesis. This is another indicator (albeit a gross one) of the large scale dynamics of the patches, they are not just forming and then varying in size in a noisy fashion.

We can similarly plot the centroid of each patch vs. frame number, see figure

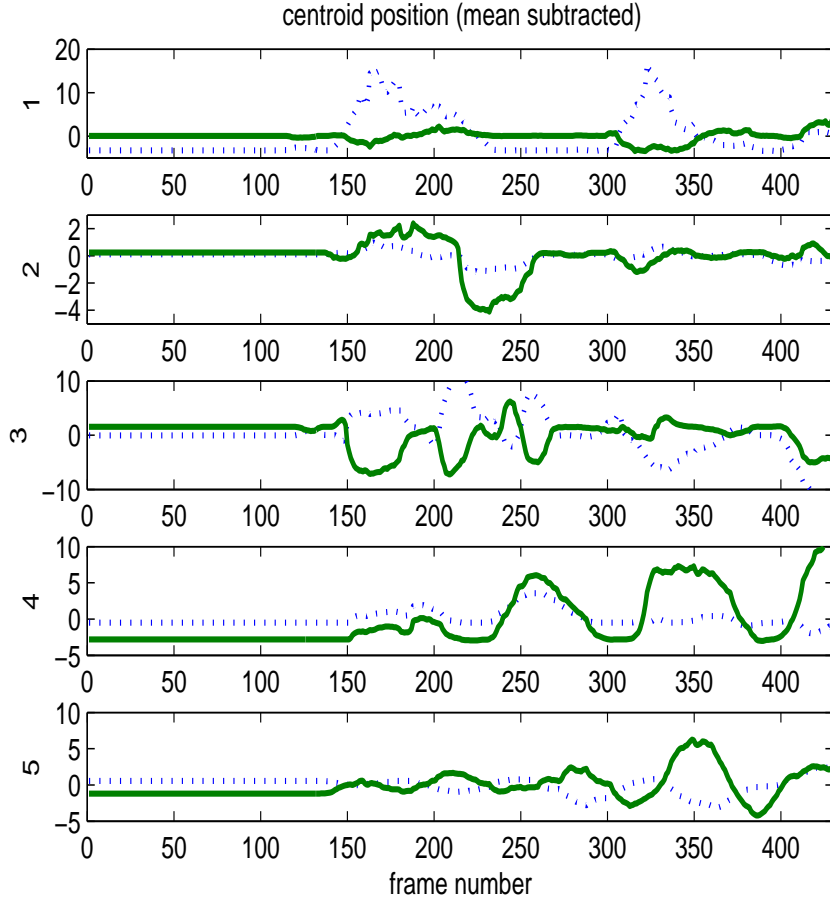


Fig. 9. Patch centroid position, mean subtracted, vs. frame number for the first five regions. X position is denoted with a dotted line, Y position with a solid line.

9. X and Y positions are plotted, with the mean of each subtracted to center the data. There will be no patch movement when the patch is not there, hence the flat spots when the patch size goes to zero. Thus the bulk motion of the patches appears correlated to the change in patch size.

4.3 Subregion Level Archetypal Analysis

A more detailed dynamics analysis requires that the shape of the patches as they evolve in time be taken into consideration. Therefore we seek to represent the patches in terms of characteristic shapes and the time evolution of these shapes. Comparing the dynamics in separate regions then becomes a question of comparing the multivariate time series formed for each region. In particular, the extent of long-range coupling between patches can be analyzed by comparing the dynamics in distinct regions over the entire leaf domain.

To reduce the patch dynamics in a single region to characteristic patterns and their time evolution we employ the method developed by Cutler and Breiman [3], known as Archetypal analysis. With this, the patch dynamics in a single region can be represented by archetypal patterns and the shift between these patterns as the frame number in the video sequence advances. Performing Archetypal analysis directly on the video images is computationally prohibitive, so the images are first compressed using Principal Component Analysis, or PCA. In the next subsection we review the necessary elements of PCA and Archetypes; see Kirby [?] for more details on the PCA applied to image data.

4.3.1 Principal Component Analysis (PCA) applied to video image data

Consider a data set $\{\mathbf{x}_i \mid i = 1, \dots, n\}$, where $\mathbf{x}_i = (x_{1i}, \dots, x_{mi})^T$. In the case of the leaf data, each \mathbf{x}_i is a column-stacked segmented video image, and m is the total number of pixel samples; for instance, a 64 by 64 image would have $m = 4096$. The ordering of the vectors is important in this application, since it holds the time information of the data set, i.e. the dynamics.

The principal components $\{\Phi_i \mid i = 1, \dots, m\}$ form a basis for the span of the data vectors. Thus for all i , there exists $\{a_{ij}\}_{j=1}^m$ such that $\mathbf{x}_i = \sum_{j=1}^m a_{ij} \Phi_j$. Furthermore, the principal component basis is optimal in the sense that the approximation $\hat{\mathbf{x}}_i = \sum_{j=1}^p a_{ij} \Phi_j$, $p < m$, of \mathbf{x}_i has minimum error. This is defined: $\varepsilon_p = \langle \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|^2 \rangle$. Here $\langle \cdot \rangle$ denotes an ensemble average, and $\|\cdot\|$ denotes the Euclidean norm. Minimizing ε_p subject to the constraint $(\Phi_i, \Phi_j) = \delta_{ij}$, where δ_{ij} is the Dirac delta, via Lagrange multipliers results in an eigenvalue problem for the Φ 's: $\mathbf{C}\Phi_j = \lambda_j \Phi_j$ where $\mathbf{C} = \langle \mathbf{x}_i \mathbf{x}_i^T \rangle$ is the ensemble averaged covariance matrix.

\mathbf{C} is an $m \times m$ symmetric non-negative matrix and consequently determines a complete set of orthogonal eigenvectors, and real, non-negative, eigenvalues. These eigenvalues can be ordered: $\lambda_1 \geq \lambda_2 \geq \lambda_3 \cdots \geq \lambda_m$ and the statistical variance of the data set in the direction of the j th eigenvector is proportional to the j th eigenvalue of \mathbf{C} . In fact, it is easily shown that: $\varepsilon_p = \sum_{j=p+1}^m \lambda_j$. The Φ 's, which appear in many contexts, are called the empirical eigenfunctions [15], coherent structures [16], etc., by the dynamical systems community.

This is a description of what is known as the “direct method” of computing the eigenvectors, and since \mathbf{C} is $m \times m$, its eigenvectors are m dimensional and each can be viewed as a video image itself. A more efficient algorithm for computing the eigen-images is the “snapshot method” [14], which capitalizes on the fact that the eigenvectors of \mathbf{C} can be more efficiently computed from a matrix formed from a space average over temporal cross-correlations; that is, an average over 4096 $n \times n$ matrices. In cases where the number of images

is much smaller than the dimension of the data set, this is obviously a much faster computation, and since \mathbf{C} is then degenerate the method provides the n eigenvectors of \mathbf{C} that have non-zero eigenvalues.

To achieve a compression of the original data set, only the first N eigenvectors will be used to approximate the data, with the corresponding N reconstruction coefficients for each frame. Typically a certain percentage of variance, which we define to be the sum of the eigenvalues of \mathbf{C} , captured by the approximation of the data is fixed. In our applications, this is between 85 and 95%. For the regions considered this corresponds to truncating to $\ell = 10$ to 15 eigenvectors. The original n m -vectors have therefore been compressed to the n ℓ -vector reconstruction coefficients $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$, where $\mathbf{a}_i = (a_{1i}, \dots, a_{\ell i})^T$. It is the reconstruction coefficients that are then passed to the Archetype algorithm.

4.3.2 Archetypal Analysis

Archetypal analysis will further reduce the data into characteristic patterns and their time evolution. The characteristic patterns are now represented by points in the truncated reconstruction coefficient space corresponding to the first 10-15 PCA eigenvectors. In order to view the patterns, they need to be reformed using the PCA eigenvectors and the specified coefficients. In this way the data are separated into a spatial component and a temporal component.

The Archetype algorithm is defined as follows. Given multivariate data (namely the PCA coefficients):

$$\{\mathbf{a}_i, \quad i = 1, \dots, n\},$$

where each \mathbf{a}_i is an ℓ -vector

$$\mathbf{a}_i = (a_{1i}, \dots, a_{\ell i})^T,$$

find p ℓ -vectors $\mathbf{z}_1, \dots, \mathbf{z}_p$ that characterize “archetypal patterns” in the data, such that data can be represented as mixtures of the archetypal patterns. The archetypal patterns are then specified: find $\mathbf{z}_1, \dots, \mathbf{z}_p$ to minimize

$$\varepsilon_p = \frac{1}{n} \sum_{i=1}^n \|\mathbf{a}_i - \hat{\mathbf{a}}_i\|^2,$$

where

$$\hat{\mathbf{a}}_i = \sum_{k=1}^p \alpha_{ik} \mathbf{z}_k, \quad \alpha_{ik} \geq 0, \quad ; \sum_k \alpha_{ik} = 1$$

$$\mathbf{z}_k = \sum_{j=1}^n \beta_{kj} \mathbf{a}_j, \quad \beta_{kj} \geq 0, \quad \sum_i \beta_{ki} = 1.$$

The $\mathbf{z}_1, \dots, \mathbf{z}_p$ are then “archetypes”, and in our example represent the archetypal reconstruction coefficient, and by extension via the PCA, archetypal pat-

terns in the data. The α_{ik} give time series that determine how much of each archetype is present in every frame.

It can be shown that for $p > 1$ the archetypes lie on the convex hull of the data set, i.e. they represent extreme data-values such that all of the data can be well-represented as convex mixtures of the archetypes. This means that archetypes are often more readily interpreted than PCA vectors, for instance. The method also guarantees that any approximation of a data point will not fall outside the space spanned by the original data set. Computing the archetypes is a nonlinear least squares problem which we solve using an alternating minimization algorithm developed by Cutler [3]. For more information on Archetypal analysis applied to dynamical systems data sets see [27] and [5].

Our program computes archetypes for 10 random starts of the optimization algorithm and chooses the one with the lowest residual sum of squares (RSS). Plotting the RSS as a function of number of archetypes gives an idea of what p works best, typically this will decrease monotonically with p , leveling off at some characteristic p for the data (see for instance figure 11). Often sets with more archetypes than this will contain near duplicates, resulting in a less efficient characterization.

As mentioned above, for data sets from video, the archetype algorithm is ungainly. The CPU time of computation increases roughly linearly with the size of the data space [4], so that even an image with dimensions 64 by 64 pixels lives in the unreasonably high dimensioned R^{4096} . Hence we use the hybrid technique outlined above: first the data is compressed using PCA together with the snap shot method; and the resulting reconstruction coefficients analyzed via archetypes.

Mixture coefficients, i.e. the α_{ik} 's, are best understood by considering the archetypes shapes and the data together, where it is evident how different points in the data are made up of mixtures of archetypal patterns. The time series created from the mixture coefficients, as the frame number is advanced, carries information about time evolution of the system.

For another example of this hybrid method applied to video data see [28].

4.3.3 Examples from the data

All 15 regions of the leaf data were analyzed via this technique, here we present the results for two different regions to illustrate the technique. The first example is region 12, which contains a single patch that forms late in the experiment, and whose evolution appears to be linked to the single patch in region 11. See figure 10 for sample video frames. The centroid is subtracted from

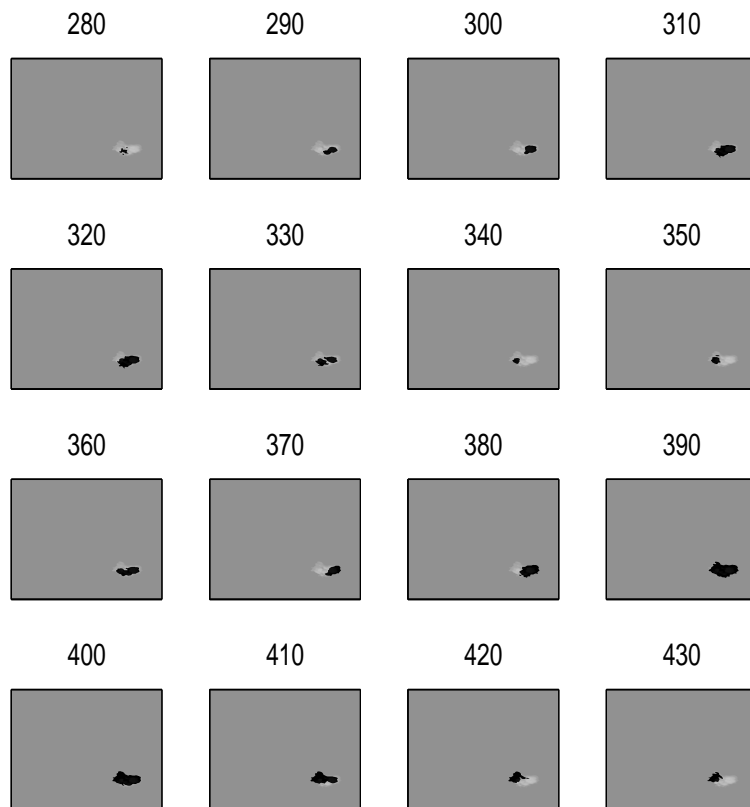


Fig. 10. Sample Frames for Region 12, Centroid Subtracted.

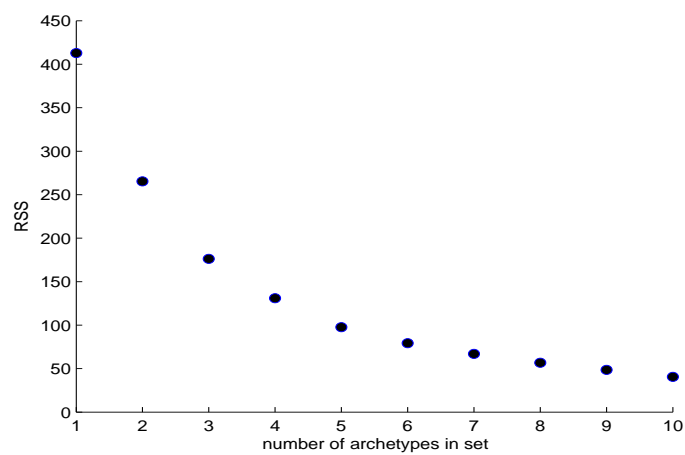


Fig. 11. Residual Sum of Squares for Region 12.

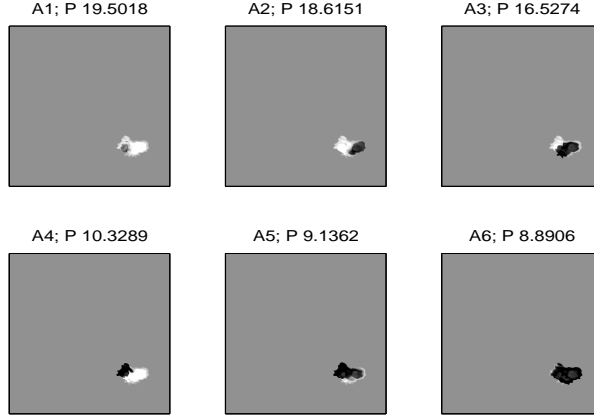


Fig. 12. Six Archetype Set for Region 12.

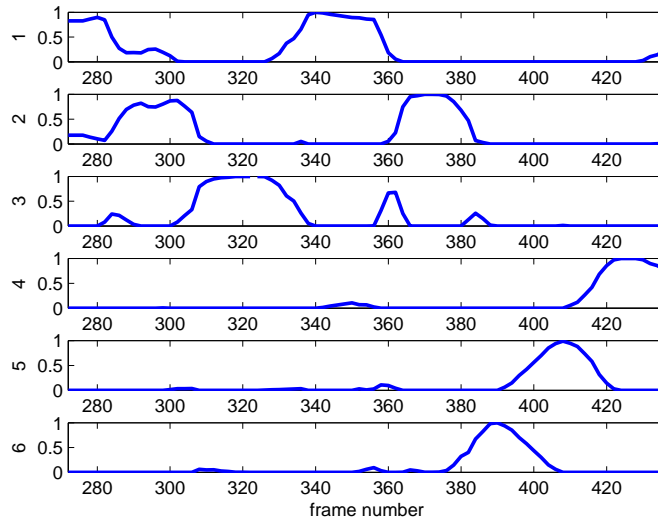


Fig. 13. Alpha Time Series for Region 12.

each vector (PCA requires this), which changes the greyscale mapping, hence the background is now grey, and the white pixels are now black. An artifact of this procedure is a white “ghost” left wherever the pixels at some point turned dark. The white patches of closing stomata are now dark patches in the image. Using PCA, we computed 16 eigenvectors for 92% of the variance, and archetypal analysis of the reconstruction coefficients in the 16 dimensional space resulted in the RSS vs. number of archetypes plotted in figure 11. This curve is typical for most of the regions, a leveling-off occurs at the 5 or 6 archetype group, so we chose to use a 6 archetype group for all the regions. This set is plotted in figure 12, ordered by the amount of “power” in each archetype, which means the sum of α_i over the frames, e.g. the sum of the value of the time series for that archetype, see figure 13. The white “ghost” appears since the centroid has been subtracted from the data before its re-

duction via PCA. It was left this way in the archetype representation since it illustrates the entire area that is ultimately occupied by the dark patch.

One fact is evident from the alpha time series: the best set of archetypes is one that has mixture coefficients that are all zero except for one “on” archetype, whose alpha will be equal to one. Archetypes are visited one after another with transitions that may involve more than the two archetypes in the transition. For instance, see frame 390 in the frame sequence for region 12 (figure 10). Archetype 6 has an alpha value there closest to 1, (see figures 12 and 13) then the pattern shifts to archetype 5 at frame 410, so that frame 400 is in between archetypes 5 and 6. The patch then shrinks to be most similar to archetype 4, as indicated by the alpha value near unity for archetype 4 at frame 420.

A slightly more complicated dynamic is seen in region 13, and we present frames from the video sequence for that region in figure 14. The patch spreads to fill the entire region, sloshes back and forth from one end to the other, and breaks into two separate patches. PCA found 14 eigenvectors captured 85% of the variance, and reconstruction coefficients for each frame in R^{14} were passed to the archetype algorithm. The resulting RSS vs. number of archetypes in a set decays much more gradually, reflecting the increase in the complexity of the dynamics.

The frames shown in figure 14 can be compared with the alpha time series (figure 17) and corresponding archetypes (figure 16). For instance, frame 280 is made up primarily of archetype 6, frame 320, archetype 4, frame 340, archetype 3, frame 360, archetype 6 again, frame 380, archetype 4, frame 420, archetype 3, while frame 400 is imperfectly captured as a mixture of 3 and 4, as the dynamics makes a transition between these two archetypes.

This sort of analysis can be carried out for all the regions, and was done so for each. The PCA cut-off ranged between 12 and 18 eigenvectors, capturing between 85% and 95% of the variance. The archetypal decomposition error leveled off between 5 and 6 archetype sets for each region, indicating a similarity in the type of the patterns within each region. Since the regions were chosen by hand, with the determining factor being coherent patch dynamics, the uniformity across regions indicates the success of this domain decomposition.

Being now convinced that the alpha time series do represent the time dependence of the patch patterns, we determine how best to analyze the relationship between the multivariate alpha time series for each region.

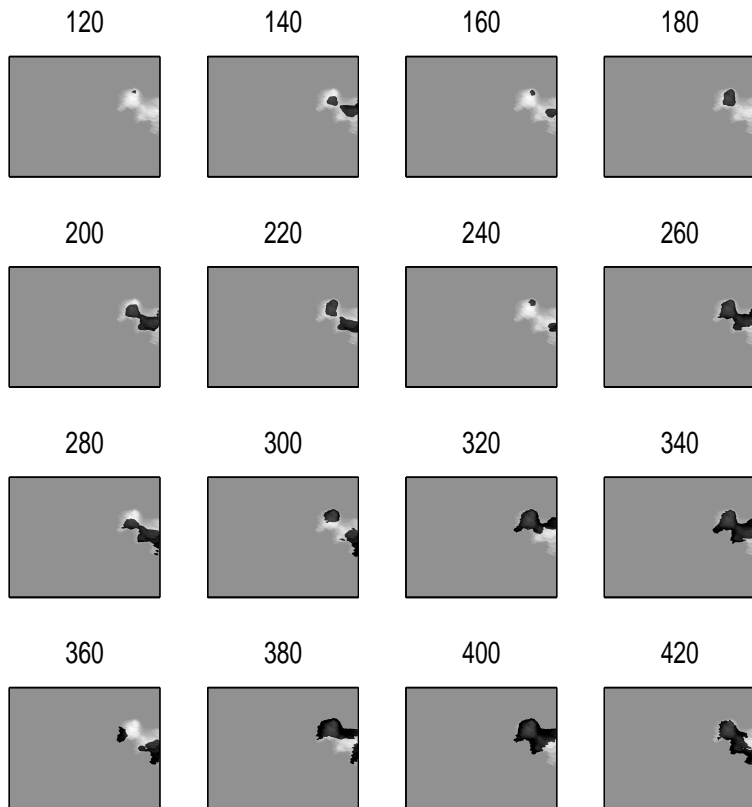


Fig. 14. Sample Frames for Region 13.

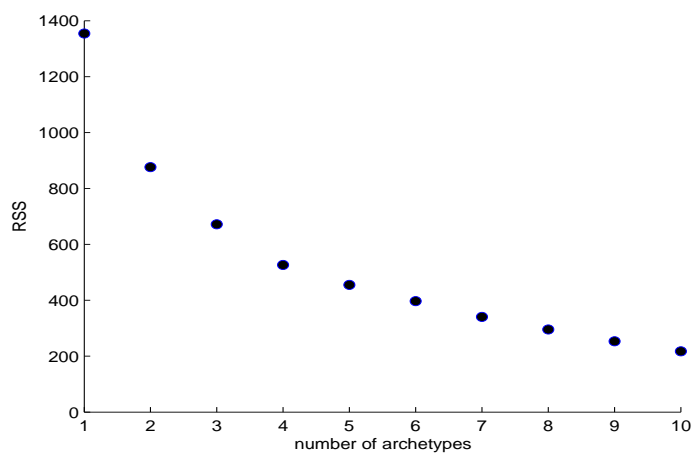


Fig. 15. Residual Sum of Squares for Region 13.

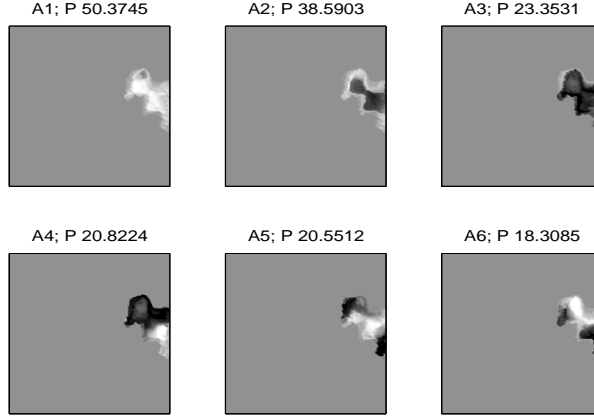


Fig. 16. Six Archetype Set for Region 13.

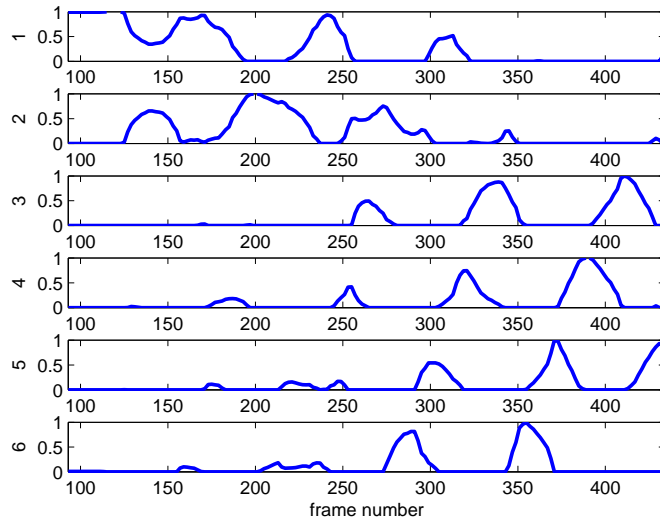


Fig. 17. Alpha Time Series for Region 13.

4.4 Time series analysis: Mutual Information

In order to compare time series of different regions we capitalize on the fact that archetypal analysis finds representative patterns in the data set, and can create a single discrete valued time series that captures the dynamics as an orbit that visits archetypal patterns in turn. To do so, for a given frame we chose the archetype number with the largest alpha and use that number in the new time series, which will then be symbol sequence on p symbols, p being the number of archetypes in the set. In figure 18 we illustrate this process for the alpha time series for region 13. This partitioning of the multivariate time series is possible because of the unique characteristics of archetypes, and naturally its effectiveness will be determined by the quality and efficiency of

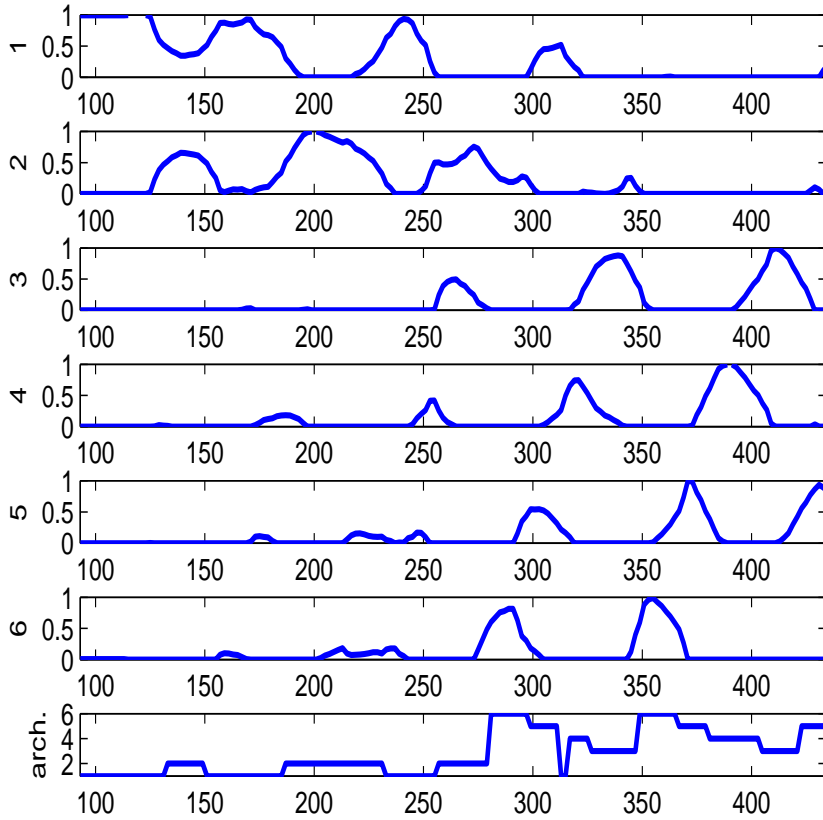


Fig. 18. Alpha time series for region 13 with 6 symbol sequence for comparison.

the archetypal representation. By this we mean that the number of archetypes must not be chosen so small as to create a large error in recreating the data, but neither should it be so small that duplicate archetypes appear in the set.

Such a symbol sequence is created for all 15 of the regions of the leaf domain, and the remaining question is how to determine statistically the relationship between the dynamics in different regions. We choose to use concepts from information theory to assess this relationship, specifically mutual information. Mutual information functions independently of actual values in the time series (unlike cross-correlation), using conditional probabilities of encountering N-tuples of symbols in the sequences being analyzed. In the following we introduce the basic idea; for a more thorough treatment see [] or [].

Let X be a discrete random variable with the alphabet \mathcal{X} (in our data this is the set of integer archetype numbers) and probability mass function $p(x) = Pr(X = x), x \in \mathcal{X}$. This random variable X can be described by its entropy

$H(X)$, or degree of uncertainty, which is defined:

$$H(X) = - \sum_{x \in \mathcal{X}} p(x) \log p(x).$$

Generally the logarithm is taken to the base 2, so that the units of entropy are bits. H is a functional of the probability distribution, not of the values taken on by the random variable, in contrast to autocorrelation or other related statistical measures. The oft-cited example is a random variable that takes on one of two values, with probability p and $1 - p$ respectively. The entropy of this distribution is 0 if $p = 1$ or 0, and rises to its maximum of 1 bit when $p = \frac{1}{2}$. Hence H will be large when the distribution is flat, and small when one value is much more likely to occur than the other.

The joint entropy of a pair of random variables, X and Y , relies on the joint distribution $p(x, y)$ and marginal distributions $p(x)$ and $p(y)$. Namely:

$$H(X, Y) = - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(x, y).$$

The relative entropy between two probability mass functions, $p(x)$ and $q(x)$ is defined

$$D(p||q) = - \sum_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)},$$

and is also known as the Kullbeck Leibler distance between $p(x)$ and $q(x)$.

Given these definitions the mutual information $I(X; Y)$ is the relative entropy between the joint distribution and the product distribution for X and Y :

$$I(X; Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log \frac{p(x, y)}{p(x)p(y)}.$$

It is the Kullbeck Leibler distance between $p(x, y)$ and $p(x)p(y)$, or $D(p(x, y)||p(x)p(y))$. A more heuristic definition is that mutual information is a measure of the amount of information one random variable has about another, that is, given knowledge of the state of one random variable, how much more predictable is the other? Note that, when X and Y are independent, $p(x, y) = p(x)p(y)$, so that $I(X; Y) = 0$.

We use mutual information to tell us something about the relationship between alpha times series of different regions. Specifically, how much information about one region's dynamics is conveyed by another region's dynamics? The mutual information (MI) of two random variables will depend in part on the joint entropy of the distributions, so to normalize the MI we divide it by the joint entropy, creating what we call a relative mutual information (not to be confused with conditional mutual information). In the following table we present these values calculated for all possible pairs of regions.

	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.47	0.29	0.33	0.30	0.30	0.30	0.17	0.21	0.23	0.23	0.28	0.25	0.15	0.14
2		0.38	0.28	0.27	0.36	0.30	0.26	0.26	0.21	0.20	0.21	0.19	0.24	0.25
3			0.19	0.27	0.33	0.33	0.19	0.27	0.20	0.17	0.23	0.19	0.13	0.20
4				0.26	0.35	0.33	0.20	0.25	0.30	0.20	0.19	0.22	0.13	0.13
5					0.36	0.24	0.17	0.23	0.27	0.30	0.45	0.42	0.17	0.16
6						0.42	0.22	0.30	0.31	0.15	0.20	0.23	0.12	0.26
7							0.26	0.29	0.31	0.14	0.17	0.19	0.12	0.30
8								0.27	0.17	0.20	0.19	0.20	0.14	0.33
9									0.30	0.15	0.19	0.23	0.17	0.26
10										0.19	0.23	0.30	0.10	0.16
11											0.43	0.26	0.16	0.14
12												0.31	0.16	0.14
13													0.12	0.15
14														0.17

These numbers are meant to be compared to one another, not to any outside norm. However, an idea of the significance of these numbers can be gained by comparing them to mutual information values from random sequences. To do this we computed the average mutual information between two time series with the same percentage of symbol repeats as seen in the data, but with a randomized order of occurrence of these symbols. This was accomplished by taking each time series in turn, randomizing it, and forming a new MI matrix. An average MI was computed for the whole matrix, and the process repeated 10 times to determine a mean and variance of the average MI computed in this way. The result is a mean of 0.0252 ± 0.0004 , compared to the mean MI value of 0.2425 ± 0.0798 for the matrix of values above.

The mean mutual information for the leaf time series is ten times that of the randomized leaf time series, increasing the significance of the implication that the dynamics in separate regions are dependent. The standard deviation cannot be compared directly, since we did not have 10 replicates of the MI matrix for the leaf data. However, the standard deviation of approximately 0.08 suggests that we can partition the values into bins of size 0.1, with reasonable significance. The average MI, 0.2425, is taken to be an intermediate value, between high and low MI. Therefore we if denote the values 0.1 – 0.2 as low MI, 0.2 – 0.3 as medium MI and 0.3 – 0.4 as high MI we can assign a

letter to each pair, H,M or L. Values above 0.4 are denoted very high or VH. The result is shown in Table 3.

	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	VH	M	H	M	H	H	L	M	M	M	M	M	L	L
2		H	M	M	H	H	M	M	M	M	M	L	M	M
3			L	M	H	H	L	M	L	L	M	L	L	L
4				M	H	H	L	M	M	M	L	M	L	L
5					H	M	L	M	M	H	VH	VH	L	L
6						VH	M	H	H	L	M	M	L	M
7							M	M	H	L	L	L	L	M
8								M	L	M	L	M	L	H
9									H	L	L	M	L	M
10										L	M	H	L	M
11											VH	M	L	L
12												H	L	L
13													L	L
14														L

As an example of the interpretation of these values, we consider the MIs for region 6, which lies in the center of the domain. It shares boundaries with regions 7,8,9,10 and 5. The MI with 7 is the largest at 0.421, next largest MI is with regions 2 (0.359) and 5 (0.358). This suggests a “super region” of the contiguous regions 2, 5, 7 and 6, all of which have high MI values with each other. Region 6 also has higher MI values with regions 3, 4 and 9, 10, all of which are either adjacent (9 and 10) or one region away from 6 (3 and 4). The rest of the MIs for 6 are medium or low.

The stand-out values for MI are the VH ratings, and those occur for regions 1 and 2, 6 and 7, 11 and 12, 5 and 12, and 5 and 13. The first three are between adjacent regions, and suggest the dynamics takes place mainly in a region formed from the union of the two. This can be confirmed by examining the video sequence, for instance regions 6 and 7 share a large patch for much of the experiment (see for instance image 400 in figure 3.) Regions 11 and 12 never share a single patch, but the high MI value reflects the synchronicity in the two isolated patches that is clear to the eye in the video. The more cryptic relationships are between 5 and 12, and 5 and 13. Since 11, 12, and 13 all have a high pairwise MI, it makes sense that if 5 has a high value with 12 it will also

have a high MI with 13. Region 5 shares a boundary with 11 and 12, though not with 13. A super-region containing 5, 11, 12 and 13 is indicated by these values, which is a large region making up the bottom right hand corner of the leaf. That the dynamics in these regions are so related is difficult to determine by an inspection of the video sequence, so here is where an objective measure such as MI has its utility.

The veins in the domain do seem to organize the behavior to some extent by setting up boundaries between the different regions. The largest veins, found at the bottom of the domain, separate the regions that have no active patches from those that do. In the case of regions 7 and 6, the MI analysis reinforces the idea that vein size is a determining factor in splitting up the patches themselves, since the vein dividing regions 7 and 6 is smaller than many of the other veins. But the relationship between 11 and 12 is strong as well, and the vein dividing the two regions is larger than most in the domain. Hence this analysis must be viewed as preliminary, no immediate conclusions can be drawn about the relationship between vein size and patch dynamics.

5 Conclusions

The stomata of respiring leaves are a physical example of a sparsely coupled network of individual elements whose dynamics are partially synchronized. The synchronization on a small scale forms patches of stomata executing similar changes in aperture. These patches form, evolve and sometimes expire, the result being a complicated pattern formation exercise in two dimensions. Extracting relevant features of the dynamics from a noisy video data set has been the goal of this project.

The numerical methods developed to create an image segmentation of the leaf video into the patches are a nontrivial extension of the segmentation approaches of [6,13] to three dimensions. In this way a segmentation that is smooth in time (as frame number is advanced) is attained without ad hoc constraints. In order to not involve the veins in the segmentation a background subtraction was performed, involving another variation of standard techniques in image processing. A detailed presentation of the segmentation method is presented in [17]. A convergence analysis of the segmentation method constitutes possible future work.

Once the segmentation was complete, analysis of the dynamics of the fluorescing patches could be executed. We analyzed both large scale (over the whole domain) and medium scale (on distinct regions) patch dynamics.

To gain an overall picture of the patch dynamics on the entire leaf domain we

computed statistics on the patch sizes. We also examined the time evolution of the number of patches and average patch size. This sort of calculation would be very difficult (if not impossible) using the original data set, while it was a straight forward procedure on the segmented data.

The spatial aspect of the patch dynamics was analyzed by dividing the domain into regions (separated by veins) where the patch activity was perceived to be coherent. The images of the patches within a region were first compressed using PCA, then decomposed into a spatial and temporal component using Archetypal Analysis. Because of the properties of archetypes, it was then possible to create a single time series for each region, one that indicated switching between archetypal patterns as the patches evolved. These integer-valued time series were then analyzed with the statistical measure of mutual information. Mutual information gives an indication of how much information about one time series is stored in another, or how much more predictable a time series is given another time series. It provides a measure of correlation (abusing the term), or synchronization, of the dynamics represented by two time series.

From the mutual information calculation we determined that while the dynamics in most neighboring regions were moderately related, the relation decayed with spatial separation of the regions, except in a few cases. Some super-regions seemed to form, in physically identifiable parts of the domain.

The analysis has three key features that make it interesting and generally useful. First, the video data is segmented using a three dimensional level set method that gave a smooth representation over the frames. Secondly, the spatially extended dynamics are resolved in an objective way into time and space components, using the hybrid PCA-Archetypal Analysis technique for image data. The properties of Archetypal analysis allow for a natural discretization of the time part of the dynamics, that is, into a symbol sequence on the labels of the archetypes themselves. This side-steps many of the issues in binning time series of continuous variables and is the third feature that is illustrated by this application.

This last point, that the dynamics of the patterns can be represented as a symbol sequence naturally with archetypes, suggests other techniques for analyzing symbol sequences - such as the construction of Markov models and hidden Markov models - to classify the complexity of the dynamics. This will require much longer video sequences, hence longer experiments. The computation of mutual information in section 4.4 is limited by the brevity of the time series, longer experiments are needed to increase the reliability of all the statistical measures.

References

- [1] Beyschlag, W. and J. Eckstein, J. Prog. Bot., vol. 60, 1998, pp. 283-298.
- [2] Buckley, T., Farquhar, G. and K. Mott, Oecologia, vol. 118, 1999, pp. 132-143.
- [3] Cutler, A. and L. Breiman, *Archetypal Analysis*, Technometrics, 1994 pp. .
- [4] Cutler, A. Ph.D. thesis.
- [5] Cutler A. and E. Stone, *Moving Archetypes* Physica D.
- [6] Vicent Caselles, Ron Kimmel and Guillermo Sapiro, *On Geodesic Active Contours*, International Journal of Computer Vision, vol. 22, no. 1, 1997, pp. 61-79.
- [7] Tony F. Chan and Selim Esedoglu, *Aspects of Total Variation Regularized L^1 Function Approximation*, SIAM Journal on Applied Mathematics, vol. 65, no. 5, 2005, pp. 1817-1837.
- [8] Tony F. Chan and Jianhong Shen, *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*, SIAM, 2005.
- [9] Tony F. Chan and Luminita A. Vese, *Active Contours Without Edges*, IEEE Transactions on Image Processing, vol. 10, no. 2, 2001, pp. 266-277.
- [10] B. Genty and S. Meyer, *Quantitative Mapping of Leaf Photosynthesis Using Chlorophyll Fluorescence Imaging*, Australian Journal of Plant Physiology”, vol. 22, no. 2, 1994, pp. 277-284.
- [11] Ron Kimmel, *Fast Edge Integration*, Geometric Level Set Methods, Ed. Stanley Osher and Nikos Paragios, Springer-Verlag, 2003, pp. 59-77.
- [12] Kirby, M. *Geometric Data Analysis*
- [13] Ron Kimmel, *Numerical Geometry of Images: Theory, Algorithms, and Applications*, Springer, 2004.
- [14] Kirby, M.
- [15] Lumley, J.
- [16] Sirovich, L.
- [17] Aaron Luttmann and Johnathan Bardsley, *A Variational Approach to Image Segmentation for Botanical Data*, Univeristy of Montana, Dept. of Math. Sci. Tech. Report No. 3, 2006.
- [18] K. W. Morton and D. F. Mayers, *Numerical Solutions of Partial Differential Equations*, Cambridge University Press, 1994.
- [19] K. A. Mott, Z. G. Cardon and J. A. Berry, *Asymmetric patchy stomatal closure for the two surfaces of Xanthium strumarium L. leaves at low humidity*, Plant, Cell and Environment, vol. 16, pp. 25-34, 1993.

- [20] D. Mumford and J. Shah, *Optimal Approximation by Piecewise Smooth Functions and Associated Variational Problems*, Communications in Pure and Applied Mathematics, vol. 42, 1989, pp. 577-585.
- [21] Stanley Osher and Ronald Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces*, Springer-Verlag, Series in Applied Mathematical Sciences, vol. 153, 2003.
- [22] Stanley Osher and James A. Sethian, *Fronts Propagating with Curvature Dependent Speed: Algorithms Based on Hamilton-Jacobi Equations*, Journal of Computational Physics, vol. 79, 1988, pp. 12-49.
- [23] David Peak, Jevin West, Susanna Messinger and Keith Mott, *Evidence for Complex, Collective Dynamics and Emergent, Distributed Computation in Plants*, Proceedings of the National Academy of Sciences, vol. 101, no. 4, 2004, pp. 918-922.
- [24] Leonid I. Rudin, Stanley Osher and Emad Fatemi, *Nonlinear Total Variation Based Noise Removal Algorithms*, Physica D, vol. 60, 1992, pp. 259-268.
- [25] James A. Sethian, *Level Set Methods and Fast Marching Methods*, Second Edition, Cambridge University Press, Cambridge Monographs on Applied and Computational Mathematics, no. 3, P.G. Ciarlet, A. Iserles, R. V. Kohn and M. H. Wright eds., 1999.
- [26] K. Siebke and E. Weis, *Assimilation image of Leaves of Glechomam Hederacea: Analysis of Non-Synchronous Stomatal Related Oscillations*, Planta, vol. 196, 1995, pp. 155-165.
- [27] Stone, E. and A. Cutler, Physica D paper.
- [28] Stone, E. *Exploring Archetypal Dynamics of Pattern Formation in Cellular Flames*, Physica D, vol. 161, 2002, pp. 163-186.
- [29] Joachim Weickert, Bart M. ter Haar Romeny, and Max A. Viergever, *Efficient and Reliable Schemes for Nonlinear Diffusion Filtering*, IEEE Transactions on Image Processing, vol. 7, 1998, pp. 398-410.