The identification of leaf venation system plays an important role in the recognition of plant species and in the analysis of plant growth. Sometimes it also helps to define the mutual localization of veins and various abnormal regions e.g. brown spots appearing in a leaf blade. In this paper detection of hydrogen peroxide (H$_2$O$_2$) performed using a cytochemical dye, 3,3-diaminobenzidine (DAB), polymerizing into a brown compound in the presence of H$_2$O$_2$ and plant peroxidases, led to appearance of spots of different size and intensity. H$_2$O$_2$ is a reactive oxygen species constitutively present in plant tissues. Its content usually increases in response to different stress factors including pathogen infection [5]. Concentration of H$_2$O$_2$, after its extraction from tissues, can be estimated using several biochemical methods, but its localization in plant organs is possible only by in situ detection of H$_2$O$_2$ e.g. after reaction with DAB. H$_2$O$_2$ can accumulate to different extend and in different regions of blade, both in veins and intervein tissues. In order to visualize spots of DAB polymer the leaf pigment chlorophyll is removed with ethanol, however the whole blade remains brownish in apple leaves rich in polyphenols and tannis. In transmitted light veins are also visible as brown objects which can sometimes overlap the spots. Because of the situation described above preliminary extraction of leaf veins, separated from the spots is needed. The algorithm to extract first and second order venation patterns would allow the quantitative description of H$_2$O$_2$ distribution in leaf blade with elimination of distortions resulting from brown coloration of vein system. The authors consider the problem to recognize primary and secondary leaf veins using a priori knowledge of pinnate venation of apple tree leaf and strict assumptions of leaf position such as petiole location and the orientation of blade.
The majority of the known methods for automatic leaf vein extraction apply to the venation systems of leaves without colour and morphological deformations and often serve for plant species recognition or the modelling of leaf growth [4, 6, 8, 11, 12]. Earlier algorithms focus on edge extraction using edge operators such as Sobel, Prewitt or Laplacian [1, 6]. More sophisticated methods apply independent component analysis (ICA) [6], artificial ant optimization [7] or the thinning of binary objects after local thresholding to recognize local vein directions [3]. Some of them also analyse vein widths to distinguish the primary from accessory veins and even use ANN classifiers after the distillation of local image features [1]. In the case of the presented apple leaves similar saturation, hue and brightness of veins and revealed by DAB staining regions of H₂O₂ accumulation disturb the process of vein extraction. Therefore to identify leaf veins the authors proposed a different approach similar to the methods of curve tracing [10, 13, 15, 18]. Because of their computational complexity deep Gaussian filtering, Hessian, Jacobian matrices calculation and all subpixel analysis were not taken into account. Robustness of the method to the artefact spots is obtained by the appropriate formulation of a cost function associated with tracking steps.

2. Material and methods

Leaves of 1-year-old apple tree (*Malus domestica* L.) were infected with *Erwinia amylovora*, bacterial pathogen causing fire blight. For H₂O₂ detection, infected as well as control leaves were stained with 3,3-diaminobenzidine (DAB) according to the method of Thordal-Christensen et al. (1997) [17]. The stained leaves were photographed using a digital camera. For good visualization of leaf venation and DAB spots localised in veins the photographs were taken in the transmitted light.

Each of the colour images obtained from digital camera had the original size of 2560×1920 pixels. To speed up software processing these images were initially reduced 3\(\frac{1}{3}\) times down to the size of 768×576 pixels corresponding to the CCIR/PAL standard.

3. The detection method of apple-tree leaf veins

3.1. Leaf blade image segmentation

The initial step of the proposed algorithm consists in the segmentation of digital leaf image into a leaf blade and its background (Fig. 1). As a result of chemical treatment and the specimen illumination described above all leaf samples have their colour hue and saturation limited in specific ranges. The background colour is intentionally selected to be most different from the leaf object in colour space. Therefore a leaf blade can be segmented by thresholding hue or saturation image components of HSV colour space or by thresholding luminance image \(I_y\) (e.g. from YIQ space) (Fig. 2) [2, 16]. The conversion from the colour space to the luminance image is provided by MATLAB function shown in equation (1).

\[
\forall (x, y) \in D \quad I_y(x, y) = rgb2gray(I_{RGB}(x, y))
\]
where:

\[ I_{\text{RGB}} \] – input colour image in RGB space,
\[ I_Y \] – output luminance image,
\[ D \] – image domain \( D = 768 \times 576 \) pixels.

**Fig. 1.** Relevant terminology to describe the parts of apple-tree leaves. Many dark spots (originally in red-brown colour) disturbing venation pattern are visible in the leaf blade.

The definitions of hue and saturation ensure the independence of segmentation result on uneven image illumination [2]. The authors selected more contrasting luminance image for further processing. It has been also proved for the series of tested \( I_Y \) images, that the relative difference of an average background luminance for all image corners is less than 10\%, while the relative difference between average foreground and background brightness is about 55\%.

**Fig. 2.** The example of apple-tree leaf with the visible after-effect of hydrogen peroxide activity in the form of darker spots: a) luminance image \( I_Y \); b) hue image \( I_H \).
Therefore the global thresholding of the image $I_Y$ has been applied using popular Otsu method from the MATLAB “Image Processing Toolbox” [9, 16] (equation (2)).

$$I_M = \sim \text{im2bw}(I_Y, \text{graythresh}(I_Y)) \quad (2)$$

where:

$I_M$ – binary leaf mask image obtained after thresholding (black in background),

$\sim$ – negation operator in M-language.

The thresholding operation is then followed by morphological openings on the binary image $I_M$ and its negation $\sim I_M$ as expressed in equation (3) [14, 16], what eliminates spurious small white objects in the background and small black gaps inside of the leaf lamina.

$$I'_M = \text{bwareaopen}(I_M, A_{\text{MIN}}), I'_M = \sim \text{bwareaopen}(\sim I_M, A_{\text{MIN}}) \quad (3)$$

where:

$I_M (\sim I_M)$ – binary image of the leaf lamina (or its negation),

$A_{\text{MIN}}$ – the minimum area value of objects not to be eliminated.

An example result of the segmentation with morphological cleaning is shown in Figure 1 in the form of leaf blade contour superimposed on the leaf luminance image $I_Y$.

### 3.2. The algorithm of vein tracking

The proposed method of vein tracking is executed stepwise and applies only local optimization in each of the steps. It does not provide a search for global fitting to the pattern of leaf veins. For each step, a tracking route with a current endpoint $P_e$ is extended by a new line segment $P_eP_{e+1}$ as shown in Figure 3. This segment is only the initial part of a selected prediction vector $P_eP_p$ with given length $R_p = |P_eP_p|$. The vector prediction in current tracing step is realized by a consecutive search for the direction, for which a given cost function is locally minimized. A priori knowledge of pinnate venation type suggests that primary and secondary veins change their angles very slowly over entire length. Then a searching domain in each step can be defined in the limited range of angles $[\theta_1, \theta_2]$ around the direction $\theta_e$ of the preceding line segment $P_{e-1}P_e$.

The proposed cost function $C$ is evaluated over the luminance image $I_Y$ as expressed in the equation below:

$$C(\theta) = k_0 + k_1 C_1(\theta) - k_2 \left( C_2(\theta, d_2) + C_3(\theta, d_3) \right) \quad (4)$$

where:

$k_j$ – (j = 0,1,2) positive constants in the formula, adjusted experimentally,

$\theta$ – the angle of prediction vector $P_eP_p$ around the previous direction $P_{e-1}P_e$,

$d_i$ – (i = 1, 2, 3) the perpendicular offset of prediction vector $P_eP_p$,

$C_i$ – (i = 1, 2, 3) i-th component of the cost function $C$ along each vector parallel to the $P_eP_p$. 

The function $C(\theta)$ exploits the average values of the brightness image $I_Y$ in the set of pixels representing the triplets of parallel prediction vectors separated by the distances $d_i$.

Hence the calculated cost values are robust to the tracking distortions like image noise and the dark spots of fire blight or folding of the leaf surface. The perpendicular offsets $d_2$, $d_3$ of associated prediction vectors (Fig. 3) should be large enough to include a tracked vein between them, although this requirement does not have critical influence on the tracking quality. The cost of current step estimated in equation (4) represents one dimensional function of tracking angle, which is expected to have one or more local minima (Fig. 4). Bearing
in mind slow angular changes of pinnate veins the authors assumed that the cost should be locally minimized at the direction $\theta_{\text{MIN}}$ where $C(\theta_{\text{MIN}})$ has its local significant minimum closest to the direction $\theta_e$ of previous tracking step.

$$\theta_{\text{MIN}} = \min_i |\Theta_{\text{smin}}(i) - \theta_e|$$

where:
- $\Theta_{\text{smin}}(i)$ – $i$-th element of the set of significant local minimum locations,
- $R_p$ – the radius of prediction sector,
- $\theta_e$ – the angle of previous step direction.

The local minimum of $C(\theta)$ is considered to be significant when it has the lowest value located between the two local maxima $c_{\text{max}}(\theta_L)$ and $c_{\text{max}}(\theta_R)$ with values exceeding this minimum by at least the significance threshold $T$ as described in equation (6). The “shallow” maxima and minima are ignored. For the purpose of significant minima detection over the entire prediction range $[\theta_1, \theta_2]$ the cost function $C(\theta)$ is extrapolated outside of this range with the constant value equal to the $C(\theta)$ global maximum found inside.

$$\Theta_{\text{smin}} = \{\theta_{\text{smin}} : (c_{\text{max}}(\theta_L) - c_{\text{min}}(\theta_{\text{smin}}) > T) \wedge (c_{\text{max}}(\theta_R) - c_{\text{min}}(\theta_{\text{smin}}) > T)\}$$

$$c_{\text{min}}(\theta_{\text{smin}}) = \min_{\theta_L < \theta < \theta_R} C(\theta)$$

$$c_{\text{max}}(\theta_L) = \max_{\theta_L < \theta < \theta_{\text{smin}}} C(\theta), \quad c_{\text{max}}(\theta_R) = \max_{\theta_{\text{smin}} < \theta < \theta_R} C(\theta)$$

where:
- $c_{\text{min}}$ – the deepest local minimum between neighbouring local maxima,
- $c_{\text{max}}$ – the local maxima of the cost function $C$ around the significant minimum $c_{\text{min}}(\theta_{\text{smin}})$,
- $\theta_L, \theta_R, \theta_{\text{smin}}$ – the cost function arguments of compared extremes,
- $\theta_L, \theta_R$ – the arguments of significant minima – previous (left) and next (right) to $C(\theta_{\text{smin}})$ if any of them exists, otherwise the respective ends of $C(\theta)$ domain,
- $T$ – significance threshold for the local minima,
- $\Theta_{\text{smin}}$ – the set of significant minimum locations.

Two example regions of secondary veins tracking are shown in Figures 5a and 5b. There can be seen that the proposed method copes well with the problem of distortions in a form of dark spots reminded after chemical treatment with diaminobenzidine.
To stop the iteration of vein tracking one or more of the following four conditions should be fulfilled:

1) the lack of significant minima of the cost function $C(\theta)$ over the all searched directions:

$$\Theta_{\text{MIN}} = \emptyset$$

(7)

2) the average brightness of the first component $C_1$ of $C$ function is greater than the average brightness of leaf blade decreased by its standard deviation,

$$C_1(\theta_{\text{MIN}}) > \text{mean}(I_Y(B)) - \text{stdev}(I_Y(B))$$

(8)

where:

- $C_1$ – the component of the cost function in equation (4),
- $B$ – the region of leaf blade,
- $\theta_{\text{MIN}}$ – the angle of optimal step direction.

3) the new selected pixel is too close to the leaf margin (closer than a certain distance $d_{\text{mrg}}$) except for the line starting in a petiole:

$$\rho(P_{e+1}, M(x, y)) < d_{\text{mrg}}$$

(9)

where:

- $\rho$ – distance between the new selected pixel $P_{e+1}$ and the leaf margin $M(x, y)$,
- $d_{\text{mrg}}$ – the lower limit of the distance $\rho$. 

Fig. 5. The examples of tracking secondary veins in the regions of luminance image distorted by red-brown spots indicating the secretion of hydrogen peroxide:

a) first case; b) second case
4) the predicted pixel of secondary vein is closer to the midrib than its predecessor:

$$\rho(P_{p}(x_{p}, y_{p}), L_{0}(x, y)) < \rho(P_{c}(x_{c}, y_{c}), L_{0}(x, y))$$  \hspace{1cm} (10)$$

where:

- $\rho$ – distance function between two arguments,
- $P_{p}, P_{c}$ – the predicted and last accepted points of vein route,
- $L_{0}$ – the route of primary vein.

When the algorithm meets the first or the second stopping conditions the tracking step size will be increased and the tracking angle range decreased up to four times. This action concerns the case of vein fading and enables to find vein’s possible continuation assuming the smaller range of tracking angles. All conditions given above are based on a priori knowledge and the observation of primary/secondary veins morphology. The example result of primary and secondary vein detection applying the considered method is shown in Figure 6.

![Fig. 6. The graphical result of primary and secondary vein tracking overlapped on the leaf blade image](image)

### 3.3. Leaf petiole and midrib start detection

It is assumed that every apple-tree leaf sample is placed horizontally inside of an image frame i.e. the vertical coordinates of its midrib pixels change very little. The leaf petiole should be located at the left boundary of image frame as shown in Figure 1. The petiole width stays small and stable over its entire length, up to the beginning of a leaf base, where the width rapidly increases with the distance from petiole (Fig. 7a). Column-wise scanning
of the leaf binary image \( I_M \) can detect in each column small vertical distances of white pixels as leaf stalk widths, while their lengths stay approximately equal each to the other. The idea of \( x \)-coordinate leaf base detection as \( x_b \) is described by equation (11).

\[
x_b = \inf \left\{ x : w(x) > k \cdot \text{med} \left( \left[ w(1), ..., w(x) \right]^T \right) \right\}
\]

(11)

where:

\( \inf() \) – the infimum operator in the set of \( I_M(x, y) \) \( x \)-coordinates, for which the above inequality condition is fulfilled,

\( w(x) \) – the petiole width along \( y \)-dimension at the \( x \)-coordinate of mask image \( I_M(x, y) \),

\( \text{med}(W) \) – the median of the petiole widths vector \( W \),

\( k \) – the constant factor from the range \([2, 5]\) adjusted experimentally.

---

The petiole identification allows to establish the starting segment \([P_{0,1}(x_1, y_1), P_{0,2}(x_2, y_2)]\) necessary for the tracking of primary vein. The coordinate \( x_1 \) represents the first column of \( I_M \), for which the petiole width \( w(x_1) \) is at least equal to 80\% of median width.

\[
\begin{align*}
x_1 &= \inf \left\{ x : w(x) > 0.8 \cdot \text{med} \left( \left[ w(1), ..., w(x) \right]^T \right) \right\} \\
x_2 &= \min \left( x_1 + \Delta x, x_b \right)
\end{align*}
\]

(12)

where:

\( \inf() \) – the infimum operator in the set of \( I_M(x, y) \) \( x \)-coordinates, for which the above inequality condition is fulfilled,

\( w(x) \) – the same as in equation (11),

\( \text{med}() \) – the median of the width vector \( [w(1), ..., w(x_b)]^T \),

\( \Delta x \) – the step in \( x \)-direction between the pixels \( P_{0,1} \) and \( P_{0,2} \).
\[ y_i = \arg\min\left( I_Y(x_i, y) \middle| I_M(x_i, y) > 0 \right), \quad i = 1, 2 \]  

(13)

where: \( x_i, y_i = (i = 1, 2) \) the horizontal and vertical coordinates of the initial midrib tracking points \( P_i \).

The changes of petiole brightness in the image \( I_Y \) along the columns \( x_1, x_2 \) are shown in Figure 7b.

### 3.4. Secondary vein origins identification

The algorithm of primary vein tracking follows only the midrib axis and ignores lateral paths. The origins of secondary veins are expected to be found on the two lines \( L_L, L_U \) parallel to the detected midrib line \( L_0 \), close to the brightness minima of \( I_Y(L_U) \) and \( I_Y(L_L) \) (Fig. 8). This approach is possible thanks to the pinnate venation type of apple-tree leaves.

The two parallel lines are defined as follows:

\[ L_0 = \{(x, y)\} \Rightarrow L_U = \{(x, y - d_0)\}, \quad L_L = \{(x, y + d_0)\} \]

(14)

where:

- \( L_0 \) – the polygonal line of a primary vein,
- \( L_L, L_U \) – lower and upper lines parallel to the primary vein (midrib),
- \( d_0 \) – vertical distance between \( L_0 \) and any of the parallel lines \( L_U, L_L \).

The location sets \( S_{s\min} = \{s_{s\min}\} \) of significant brightness minima are found according to the rule in equation (6) but are expressed as the parameters of each curve \( L_U \) and \( L_L \) instead of prediction angles. Every tested location \( s_{s\min} \) is then expanded to its neighbourhood \( N(s_{s\min}, R_L) \) with a certain radius \( R_L \), because the origin of a secondary vein can be obscured by an accidental dark spot but it should always stay close to the minimum \( s_{s\min} \).

---

![Fig. 8](image_url)

**Fig. 8.** The example region of a leaf image with the detected axis of primary vein, associated parallel lines \( L_U, L_L \) and marked initial steps of secondary vein tracking (a) the explanation of the first step of vein tracking (b) \( L_0 \) – primary vein axis, \( L_i, L_j \) – secondary veins, \( P_{i,1}, P_{j,1} \) – starting pixels, \( \theta_i \) – starting angle.
Then for any starting pixel $P_{i,1}$ around the significant minimum $P_{s\text{min}}$ (Fig. 8b):

$$s_i \in \{N(s_{s\text{min}}, R_L)\}, \quad P_{i,1} \in \{N(P_{s\text{min}}, R_L)\}$$

(15)

where:

$N(s_{s\text{min}}, R_L)$ – the neighbourhood of any significant minimum parameter $s_{s\text{min}}$ with the radius $R_L$ along the line $L_1(L_U)$,

$N(P_{s\text{min}}, R_L)$ – the neighbourhood of any significant minimum pixel $P_{s\text{min}}$ with the radius $R_L$ along the line $L_1(L_U)$.

The tracking of secondary veins starts with the vector $P_{i,1}P_{i,2}$ which minimizes the cost function given in equation (4). Typical tracking step includes scanning the symmetric range of angles ($\theta_2 - \theta_c = \theta_c - \theta_1$ in Fig. 3) around the previous direction. The first tracking step with no predecessor tests the vector $P_{i,1}P_{i,2}$ with an angle in the fixed ranges $[\theta_{L1}, \theta_{L2}]$ or $[\theta_{U1}, \theta_{U2}]$. This angle is measured between a new direction and the midrib local direction at the pixel $P_{i,1}$. The ranges of tested angles are based on a priori knowledge about pinnate venation of the tested apple-tree leaves. The finally selected start direction $P_{i,1}P_{i,2}^*$ (equation (16)) minimizes the cost function $C$ similar to that in equation (4), but dependent on two arguments ($s_i$, $\theta_i$) without initial preference of any angular location.

$$P_{i,1}P_{i,2}^* = \arg \min_{s_i, \theta_i} C(P_{i,1}P_{i,2}(s_i, \theta_i)), \quad P_{i,1} = P_{i,1}(s_i), \quad P_{i,2} = P_{i,2}(s_i, \theta_i)$$

(16)

where:

$C$ – the cost function of tracking, evaluated for the distance $P_{i,1}P_{i,2}$,

$s_i$ – the line position of the pixel $P_{i,1}$ in the neighbourhood of a significant minimum,

$\theta_i$ – the angle at the pixel $P_{i,1}$, between $P_{i,1}P_{i,2}$ vector and the tangent of primary vein.

### 4. Experimental results and conclusions

The proposed method is still under development. So far it has been tested for a series of 5 apple-tree leaf images acquired under the conditions described in chapter 2. Before software processing the images were reduced 3 times in size, down to 768x576 pixels. The method has been executed in MATLAB 2008 environment using a PC with dual core processor Intel Core (TM)2 Duo T5750 2 GHz, 4 GB RAM and the operating system Windows Vista Home Premium. Example execution times for the leaf image shown in Figure 6 are presented in Table 1.

<table>
<thead>
<tr>
<th>Task Execution Time (s)</th>
<th>Leaf Image Execution Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The algorithm of vein tracking has been written in M-language as a series of functions with possible code vectorization to get quicker and cleaner solution. After the conversion of algorithm functions into C++ MEX files the times of tasks execution given in Table 1 can be shortened up to 5–10 times.
Additionally, at the step of vein tracking the method can be speed up by the following actions:

- the initial storage in arrays of line segment pixels corresponding to the prediction vectors $R_p(\theta)$ (Fig. 3) for all possible directions in the image raster, the pixels $(x, y)$ of each segment are mapped using Bresenham algorithm between $P_0 = (0, 0)$ and $P_1(\theta)$, where $\theta \in [0, 2\pi]$,
- the tabulation of the two distance vectors $d_2, d_3$ (Fig. 3) for each prediction vector.

The explicitly stored line segment pixels for prediction directions $R_p(\theta)$ can be quickly changed into the points around the current point $P_c$ using translation by the vector $P_0P_c$.

The considered algorithm has several parameters adjusted experimentally to the series of analyzed images. They may need to be changed when leaf staining or the conditions of image acquisition (illumination, background etc.) will be different. The most important parameters assumed in the algorithm are as follows:

- prediction vector length $R_p = 30$ px (Fig. 3),
- step vector length $R_c = 5$ px (Fig. 3),
- the symmetrical range of prediction angles $\theta_2 - \theta_e = \theta_e - \theta_1 = 40^\circ$ (Fig. 3),
- the distance of associated prediction vectors to the central vector $P_cP_p$: $d_2 = d_3 = 5$ px (Fig. 3),
- the distance between the primary vein $L_0$ and any of the associated parallel lines $L_L, L_U$ $d_0 = 15$ px (Fig. 8),
- the absolute ranges of the first step prediction angles $[\theta_{L_1}, \theta_{L_2}] = [15^\circ, 90^\circ]$ and $[\theta_{U_1}, \theta_{U_2}] = [-90^\circ, -15^\circ]$,  
- the minimum distance of a predicted vein point to the leaf margin $d_{mg} = 15$ px (equation (9)).
The method has not been designed for common purposes like the modelling and visualisation of venation patterns or the recognition of leaf morphology. It does not use Voronoi diagram, artificial ants or any global optimization [7, 12]. Therefore it is hard to compare it with one of the common methods. It is logically simple and relatively fast. According to the method primary and secondary veins of apple-tree leaves are tracked by the selection in each step a new direction minimizing the cost function given in equation (4). The tracking will be finished when at least one of the stopping conditions included in equations (7)–(10) is fulfilled.

The method is intended to be part of a larger project identifying brown fire blight spots located between, touching or overlapping the veins. This project is also expected to determine the locations of these spots against leaf venation system. The graphic results of vein tracking obtained for a series of tested images similar to this in Figure 6 have been accepted by biologists using visual inspection method. In the future the polygonal vein paths will be smoothed with cubic splines and the widths of primary/secondary veins will be estimated.

References