DETECTION AND CHARACTERIZATION OF WEB VIBRATIONS BY ARTIFICIAL VISION

by

**C. Doignon and D. Knittel**
University of Strasbourg
FRANCE

ABSTRACT

In this paper, we address the problem of the detection of transverse web vibrations by means of a digital camera and a laser dots pattern device. For this purpose, we briefly present some video processings which yield an accurate location of the projection of laser dots in the image. However, the main contribution of this work is in the new approach developed for estimating, on-line, transverse web vibration frequencies with a geometrical method and without contact. To do so, some techniques commonly used in the computer vision field are extended to some classes of non-rigid objects suitable to account for a part of an elastic web. Finally, we compare the estimated frequencies with those provided by a simple vibrating elastic string model, in motion.

NOMENCLATURE

$\mathbf{A}$, $\underline{a}$, $\underline{a}^T$, a	matrix, column vector, row vector, scalar or a geometric point
Π	perspective projection matrix,
R_c, C_c	camera reference frame, camera origin,
R_L, C_L	laser reference frame, laser source center,
$\mathbf{K}$	matrix of intrinsic parameters of {frame grabber+camera+lens},
$\mathbf{K}_L$	matrix of intrinsic parameters of the laser device,
$\mathbf{R}, \underline{t}$	Rotation and position of a projected laser spot w.r.t. camera frame,
$\mathbf{R}_{LC}, \underline{t}_{LC}$	Rotation and position of the laser frame w.r.t. camera frame, (expressed in the laser frame R_L),
λ	coordinate of a projected laser spot onto the web w.r.t. an object frame,
$d_\perp$	(orthogonal) distance between the camera origin and a line which hold a set of collinear projected laser spots,

$\underline{r}_1$ orientation of a set of collinear projected laser spots (unit vector $\underline{r}_1 = [1\ 0\ 0]\ \mathbf{R}$),

α interbeam angle of the laser device,

f_L focal length of the laser (zoom factor),

V web velocity,

T web tension,

L length of span between the tractor and the winder,

t, ϵ some thresholds.
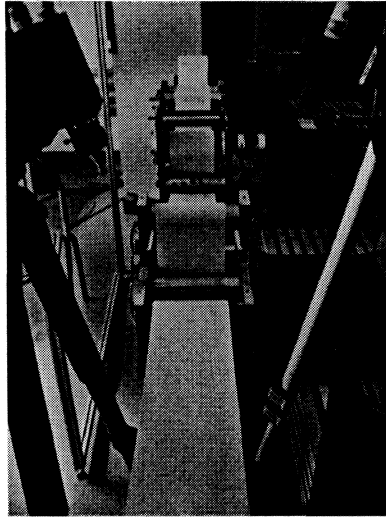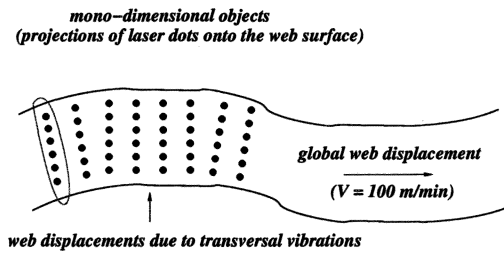
(a)   (b)

Figure 1: - (a) : while winding process, the laser dots pattern is projected onto the web surface which in turn is captured by a camera - (b) : web displacements in the direction perpendicular to the web surface can be estimated.

INTRODUCTION

A common objective in the web transport industry is to increase as much as possible the velocity while controlling tension of the web. However, some disturbances drastically limit this velocity like time-varying eccentricity and non-circularity of the roll, web sliding, temperature variations, variations in motor torque,... which are some limiting factors of performances. Since there exists a coupling introduced by elastic property of the web, disturbances are transmitted to the web tension, resulting in a web break or fold ([9],[15]). In an attempt to reduce these harmful effects, recent works on modelization, control or computer vision for web handling applications are promising ([8],[11]) and the use of digital images for supervising the quality of winding process is being emerging ([6],[4]).

The underlying application of this work is to estimate web vibrations properties of winding systems, on-line, since we have been motivated by the important economical impact of web vibrations phenomena. In this context, we present a new method for detecting transverse web vibrations by means of a vision system which consists on a digital camera and a dot

matrix-patterned laser device. The main contribution of this work is to provide a technique for simultaneously estimating web transverse displacements and the geometric transformation between the web surface (at each optical marker location, corresponding to a projected laser dot) and the camera reference frame provided that a *in situ* calibration procedure has been previously done [3]. As a result, during a sequence of images, vibrating properties (frequency and magnitude) can be estimated without unsuitable contact with the web.

The paper is organized as follows. In the next section, we give some pre-requisites about the geometry involved by the whole vision system. The extraction of visual informations, that is the location of the projection of laser dots in the image is explained in the third section, while the recovery procedure for the transformation between a cross-section of the web and the camera frame is given in section four. Experimental results and a comparison study of the estimated frequencies versus theorytical ones are discussed in the final part.

THE IMAGING MODEL

Considering a pin-hole camera model [5] (see figure 2), a point $P = (x, y, z, 1)^T$ in the 3-D projective space $\mathcal{P}^3$, expressed in the camera reference frame (R_c) and its corresponding projection $Q = w(u, v, 1)^T$ onto the image plane I (a 2-D projective space $\mathcal{P}^2$) are related by a projective transformation from $\mathcal{P}^3$ to $\mathcal{P}^2$, called a *perspective* projection, and represented by a (3×4) real matrix $\mathbf{\Pi}$:

$$Q \equiv \mathbf{\Pi} \, P \qquad (\equiv \text{ means up to a scaling factor}) \qquad (1)$$

or:

$$w \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{K} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\Pi} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \qquad (2)$$

where $\mathbf{K}$ is the (3×3) matrix of the camera parameters. These intrinsic parameters include the main geometrical characteristics of the whole vision system, that is the focal length f of the lens, sizes of frame buffer (pixels), sizes and numbers of CCD camera cells and possibly a tiny offset $(u_0, v_0)^T$ between the optical axis and the center of the CCD matrix. In other words, the matrix $\mathbf{K}$ represents an affine transformation which relates the projection point $p = (x/z, y/z, f)^T$ in the camera reference frame (R_c) (mm unit) and the projection Q in the image plane (pixel unit).

$$\mathbf{K} = \begin{bmatrix} \underline{k}_1^T \\ \underline{k}_2^T \\ \underline{k}_3^T \end{bmatrix} = \begin{bmatrix} k_x & k_{xy} & u_0 \\ 0 & k_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \approx \begin{bmatrix} k_x & 0 & 0 \\ 0 & k_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad (3)$$

$\mathbf{K}$ is an upper triangular matrix and from this expression and eq. (2), it is clear that $w = z$. If the image reference frame origin is chosen at the center of the frame buffer, the off-diagonal components of $\mathbf{K}$ are very small with respect to those in the diagonal; k_{xy} is usually setting to zero since it is accounting for a small angle between u and v axes in the image. This shearing parameter is rather not significant with recent vision sensors.
On the other hand, given a reference frame (R_o) attached to a rigid object, it's well-known that a point P with homogeneous coordinates expressed in that frame and also expressed
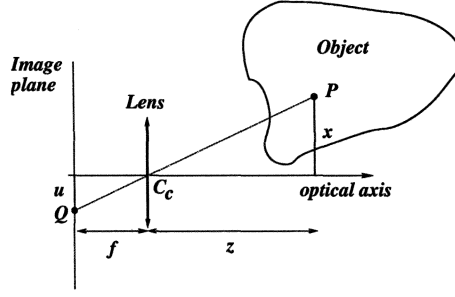
Figure 2: The stenope (or pin-hole) is a geometrical model commonly used for the representation of the perspective projection induced by the CCD camera. It's also used here to model the laser device. The projection of point P onto the image plane, Q, is passing through the lens center C_c.

in the camera reference frame (R_c) are related by:

$$P_{/(R_c)} = \begin{bmatrix} \mathbf{R} & \underline{t} \\ \underline{0}^T & 1 \end{bmatrix} P_{/(R_o)} \tag{4}$$

where $\mathbf{R} = (\underline{r}_1, \underline{r}_2, \underline{r}_3)$ is a rotation matrix and $\underline{t}$ is the position vector of the origin of (R_o) with respect to the camera frame. These parameters are called *extrinsic* parameters since they are representing the viewpoint. If the point P always belongs to a straight line Δ (an object line) with direction vector $\underline{x}_o$, the relationship between the projection Q in the image plane and $P = (\lambda, 0, 0, 1)^T$ is reduced to:

$$z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} \underline{r}_1 & \underline{t} \end{bmatrix} \begin{bmatrix} \lambda \\ 1 \end{bmatrix} \tag{5}$$

Since a straight line in space is transformed into a straight line under a projective transformation (without distortion and excepted some degenerate cases), then the point $Q = \Pi\ (P)$ belongs to a straight line $\delta = \Pi(\Delta)$. As a consequence, it's obvious that a set of collinear points P_i are projected onto the image plane as a set of collinear points Q_i. We are considering that the laser dots pattern is projected in that way either onto the web surface and either in turn onto the image plane of the camera.

A geometric model also based on perspective projection can be built for the laser pattern device. The interbeam angle α is generally known with a good accuray (provided by the manufacturer) or may be identified with an off-line calibration procedure. A laser plane is defined as a plane perpendicular to the optical laser axis and a "fixed" virtual image is then built with intersections of all laser beams and the laser plane at a given distance f_L, a focal length) from the laser source center C_L (see figure 3). In a way similar to the camera device, a matrix $\mathbf{K}_L$ of intrinsic parameters is defined for the laser pattern device:

$$\mathbf{K}_L = \begin{bmatrix} \underline{k}_{L1}^T \\ \underline{k}_{L2}^T \\ \underline{k}_{L3}^T \end{bmatrix} = \begin{bmatrix} \tan(\alpha)\, f_L & 0 & 0 \\ 0 & \tan(\alpha)\, f_L & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{6}$$
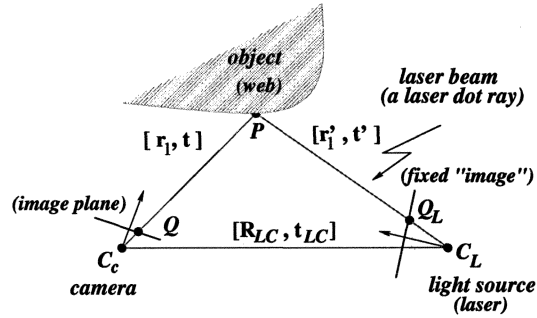
Figure 3: - The triangulation with a set of collimated laser beams and a camera. Points P, Q, C_L (laser center) and C_c (camera center) are coplanar (*epipolar constraint*).

A relationship between the projection $Q_L = w_L(u_L, v_L, 1)^T$ in the "fixed" virtual plane and $P = (\lambda, 0, 0, 1)^T$ onto the web is expressed as $(w_L = z_L)$:

$$z_L \begin{bmatrix} u_L \\ v_L \\ 1 \end{bmatrix} = \mathbf{K}_L \begin{bmatrix} \underline{r}_1' & \underline{t}' \end{bmatrix} \begin{bmatrix} \lambda \\ 1 \end{bmatrix} \tag{7}$$

In the proposed method, several assumptions have been held :

- the laser dots matrix is considered as a "fixed" virtual image, since all laser beams are issued from a common center according to a perspective back-projection [13],

- the $(n \times n)$ projected laser dots onto the web surface (a grid) are splitted into n sets of n collinear laser dots,

- each set of collinear projected laser dots is to be understood as a *mono-dimensional* object in the direction perpendicular to the web transport (see Figure 1.b). This cross-section of the web is assumed to be rigid. This seems to be a strong assumption, but throughout this paper, we will discuss about the possibility to relax this hypothesis,

- we use the linear formulation of a vibrating elastic string, in motion, stretched under a longitudinal tension between two rolls. The displacement induced by vibrations takes place entirely in one plane (the string is assumed to be perfectly flexible, that is it can transmit force only in the direction of its length). This is a quite satisfactory model as long as the web width/length ratio (between two rolls) is still small.

EXTRACTION OF VISUAL INFORMATIONS

This section is dedicated to the extraction of informations from image data in order to locate the projected laser dots in the image. The motivation for using structured lighting is based on the expectation of precise detection of the projected light patterns in the acquired images. In practice, the precise detection of simple geometric patterns is of prime importance because they determine the achievable accuracy of the 3D reconstruction. That's why implemented algorithms for image segmentation should lead to very accurate results. Here, we give a brief description of these algorithms, most of them are usual in image analysis.

Once the image is grabbed and digitized, an histogram equalization is performed to enhance contrast intensity. Contours of the digital image are then detected by convolving the intensity signal with a discrete approximation of the first derivative of a 2D gaussian at each pixel location (typically for small neighbourhood : the Sobel's filter [2]). The resulting image (gradient magnitude, see figure 4.b) is then binarized by means of an histogram-based thresholding. Since CCD camera cells are satured inside an area surrounding a laser spot, this histogram always contains two main peaks (see figure 4.c) so that we implemented the technique due to Otsu in 1979 [12] for *bi-modal* histograms. The Otsu's method is based on the minimization of the within inter-groupe variance (a peak is understood as a pixels group). The resulting image is shown in figure 4.d. An erosion (a very simple mathematical morphology operation) is applied for reducing noise and for removing some outliers (many of them are due to laser beam echos).

A list of image laser centers is built with a run-length coding algorithm [2] (also called *contours tracking* or *contours following*) dedicated to closed contours as it appear in a binary image (a connected component algorithm may be equally used for this purpose). As one can notice, by observing figures 4.d, all outliers have not been yet removed.

The last segmentation step is a classification procedure based on points labeling : the objective is to build the n sets of collinear image laser centers. In other words, the goal here is to find the n equivalence classes inside the full list of image laser centers accounting for the following relation "collinear laser spots projected onto the same cross-section of the web". To achieve this classification, the algorithm must take into account several missing spots as well as the presence of several outliers. The classification algorithm we proposed is organized as follows :

Feature point classification

1. browse the list of image laser centers and choose the first unlabeled image laser center (= a point) Q_{0k}. This point is the first (putative) point of a new equivalence class C_k (or may be an outlier) and is labeled as C_k *a priori*.

2. among unlabeled points, find the closest point, Q_{1k}, according to the vertical direction. This point is labeled as C_k if the euclidean distance between these two points is less than a threshold t, otherwise Q_{0k} is labeled as an outlier and the algorithm returns to step 1.

3. compute the straight line L_{mk} parameters (slope and distance from image origin) holding these $m = 2$ points inside the class C_k.

4. among unlabeled points, find the closest point, Q_{mk}, in the direction defined by L_{mk} up to a threshold ϵ for the orthogonal distance $d(L_{mk}, Q_{mk})$. Q_{mk} is labeled as C_k if $d \leq \epsilon$, then m is incremented.

5. repeat steps 3 and 4 until no point is added to this class, then go to step 6.

6. k is incremented. The algorithm returns to step 1, unless all points are labeled. $n \leftarrow k$ is the number of equivalence classes, that is the number of sets of collinear image laser spot centers.

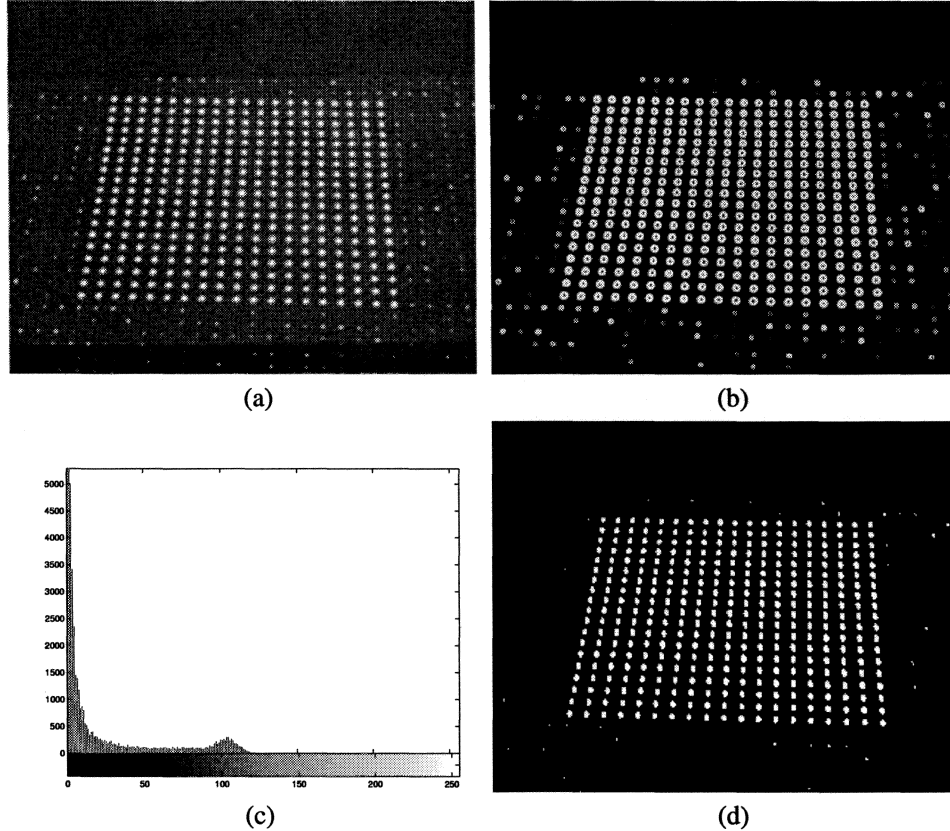(a)   (b)

(c)   (d)

Figure 4: - Results of the image segmentation for projected laser spots location.
(a) : original image. The laser dots are projected on the web surface. - (b) : contours
detection with Sobel filter (image of gradient magnitude) - (c) : histogram of image b.
which contains two peaks (the left narrow peak is the background, the right spread peak is
the group of edges pixels) - (d) : binary image after thresholding.

Figure 5 shows the result of this classification (colours and patterns (circle,cross and star)
are mixed to overcome the lake of colours available with Matlab software). Image laser
centers with identical colour and pattern are membership of the same equivalence class.
Finally, since the point order is a projective invariant, all class members (an array) are
sorted in a way that adjacent index array correspond to adjacent point (from up to bottom
ordinate, for instance). This is suitable for the further matching process which will find
correspondences between image laser centers (acquired with the camera) and laser pencils
in the "fixed" virtual image as a geometric representation for the laser beams.

ESTIMATION OF THE WEB-CAMERA DISTANCE

The motivation for using structured lighting is to offer a new measurement technique with-
out contact and is also based on the expectation of precise detection of the projected light
patterns in the acquired images. Once this is done, we are presenting the 3D reconstruction
in this section. Reconstruction, in space, for 1D rigid objects (each set of collinear laser

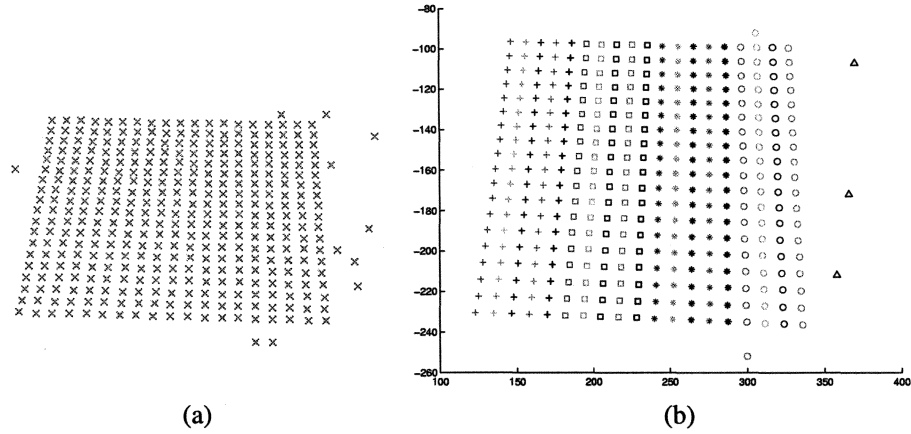(a)   (b)

Figure 5: - Results of the classification. - (a) : laser spots location (crosses) after an erosion and centroid computation - (b) : laser spots labeling. Spot centers displayed with the same colour and pattern (+,□,o,*) belong to the same line (outliers are displayed with a '△').

dots is understood as a 1D object), is not equivalent to the reconstruction of lines in space, since the relative positions of laser dots inside a line and images of these relative positions provide more informations than the only straight line and its image. Reconstruction of lines in space requires three images (that is three different viewpoints) whereas reconstruction of 1D object (with at least three collinear points) is possible with only two views as we will show below. These two views consist on the image plane of the camera and a laser plane. The camera is also assumed to be calibrated with a procedure identical to [3], this means that components of the $\mathbf{K}$ matrix are known and constant. Furthermore, the relative attitude (rotation $\mathbf{R}_{LC}$ and position $\underline{t}_{LC}$) of a laser frame, (R_L), attached to the laser image, with respect to the camera frame, (R_c), is assumed to be identified. The whole vision system can then be seen as a *stereo-like* vision system (but one of the two views is a fixed one) as illustrated in figure 3.

The above mentioned calibration procedure is only needed for estimating web vibrations magnitudes, but it's not necessary for estimating the (possibly time-variant) web vibrations frequency since a projective transformation (a 1-D homography in that case) can always be determined with an uncalibrated camera and a laser device. Our technique is based on the estimation of the geometric transformations (the unit vector $\underline{r}_1$ (2 dof) and the position vector $\underline{t}$ (3 dof) between a set of collinear projected laser dots and the camera reference frame). In a second step, the web transverse displacements can be recovered at any point in the workspace bounded by the optical markers.

The imaging model previously represented by equations (5) is now further developed on each component:

$$
\begin{aligned}
z\,u &= \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \mathbf{K}\,(\lambda\,\underline{r}_1 + \underline{t}) = \underline{k}_1^T\,(\lambda\,\underline{r}_1 + \underline{t}) \\
z\,v &= \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \mathbf{K}\,(\lambda\,\underline{r}_1 + \underline{t}) = \underline{k}_2^T\,(\lambda\,\underline{r}_1 + \underline{t}) \\
z &= \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \mathbf{K}\,(\lambda\,\underline{r}_1 + \underline{t}) = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} (\lambda\,\underline{r}_1 + \underline{t})
\end{aligned}
\tag{8}
$$

By substituting this latter expression for z in the first two equations, coordinates $(\lambda, 1)^T$ each laser dot and coordinates $(u, v, 1)^T$ of its projection in the image plane of the camera provide two equations:

$$(\underline{k}_1^T - [\begin{array}{ccc} 0 & 0 & u \end{array}])(\lambda \underline{r}_1 + \underline{t}) = 0$$
$$(\underline{k}_2^T - [\begin{array}{ccc} 0 & 0 & v \end{array}])(\lambda \underline{r}_1 + \underline{t}) = 0 \tag{9}$$

In a similar way, a relationship occurs for each laser dot and its projection $(u_L, v_L, 1)^T$ onto the "fixed" virtual image of the laser device:

$$(\underline{k}_{L1}^T - [\begin{array}{ccc} 0 & 0 & u_L \end{array}])(\lambda \underline{r}_1' + \underline{t}') = 0$$
$$(\underline{k}_{2L}^T - [\begin{array}{ccc} 0 & 0 & v_L \end{array}])(\lambda \underline{r}_1' + \underline{t}') = 0 \tag{10}$$

Since the whole vision system is calibrated, we have:

$$\underline{r}_1' = \mathbf{R}_{LC}\, \underline{r}_1 \quad , \quad \underline{t}' = \underline{t}_{LC} + \mathbf{R}_{LC}\, \underline{t} \tag{11}$$

then:

$$(\underline{k}_{L1}^T - [\begin{array}{ccc} 0 & 0 & u_L \end{array}])\, \mathbf{R}_{LC}\, (\lambda \underline{r}_1 + \underline{t}) = -(\underline{k}_{L1}^T - [\begin{array}{ccc} 0 & 0 & u_L \end{array}])\, \underline{t}_{LC}$$
$$(\underline{k}_{L2}^T - [\begin{array}{ccc} 0 & 0 & v_L \end{array}])\, \mathbf{R}_{LC}\, (\lambda \underline{r}_1 + \underline{t}) = -(\underline{k}_{L2}^T - [\begin{array}{ccc} 0 & 0 & v_L \end{array}])\, \underline{t}_{LC} \tag{12}$$

Each correspondence $(Q_{Li} \leftrightarrow P_i \leftrightarrow Q_i)$ provides four relations (eq. (9) and (12)) where the unknowns are $\underline{r}_1$, $\underline{t}$, and λ_i and are gathered in a three-dimensional vector $\underline{b}_i = \lambda_i\, \underline{r}_1 + \underline{t}$. Points C_c (the camera center), Q_i, P_i, Q_{Li} and C_L (laser source center) lie on a common plane (in absence of noise) [10], called the epipolar plane. As a consequence, only three of the four previous equations are linearly independant. This is equivalent to the following system :

$$\underbrace{\begin{bmatrix} (\underline{k}_1^T - [\begin{array}{ccc} 0 & 0 & u_i \end{array}]) \\ (\underline{k}_2^T - [\begin{array}{ccc} 0 & 0 & v_i \end{array}]) \\ (\underline{k}_{L1}^T - [\begin{array}{ccc} 0 & 0 & u_{Li} \end{array}])\, \mathbf{R}_{LC} \\ (\underline{k}_{L2}^T - [\begin{array}{ccc} 0 & 0 & v_{Li} \end{array}])\, \mathbf{R}_{LC} \end{bmatrix}}_{\mathbf{C}_i}\, \underline{b}_i = -\begin{bmatrix} 0 \\ 0 \\ (\underline{k}_{L1}^T - [\begin{array}{ccc} 0 & 0 & u_{Li} \end{array}])\, \underline{t}_{LC} \\ (\underline{k}_{L2}^T - [\begin{array}{ccc} 0 & 0 & v_{Li} \end{array}])\, \underline{t}_{LC} \end{bmatrix} \tag{13}$$

where $\mathbf{C}_i$ is a rank-3 (4×3) real matrix. For each correspondence i, this system still has three linearly independent equations, however, because of numerical tractability, it's possible that including all four equations may give a better-conditioned set of equations. A solution is then given by computing the pseudo-inverse of $\mathbf{C}_i$. Thus, solving for all $\underline{b}_i$'s within a set of m collinear points leads to m equations :

$$\lambda_1\, \underline{r}_1 + \underline{t} = \underline{b}_1$$
$$\lambda_2\, \underline{r}_1 + \underline{t} = \underline{b}_2$$
$$\cdots$$
$$\lambda_m\, \underline{r}_1 + \underline{t} = \underline{b}_m \tag{14}$$

The cross product ($\times$) with the vector $\underline{r}_1$ in both side eliminates the λ_i's. Finally, once the vector $\underline{r}_1$ is estimated (see below), the orthogonal distance from the line which holds the set of m collinear projected laser spots and the camera origin is given by :

$$d_\perp \ = \ \| \ \underline{t} \ \times \ \underline{r}_1 \ \| \ = \ \frac{1}{m} \ \| \ (\sum_{i=1}^{m} \underline{b}_i) \ \times \ \underline{r}_1 \ \| \tag{15}$$

The orientation of this line is given by the unit vector $\underline{r}_1$. It can be computed by substracting the previous equations (14) each other in order to eliminate the unknown vector $\underline{t}$. Let us note $l_{ij} = -1/(\lambda_j - \lambda_i)$, then we have :

$$\begin{aligned}
\underline{r}_1 \ + \ l_{12} \ (\underline{b}_2 - \underline{b}_1) \ &= \ \underline{0} \\
\underline{r}_1 \ + \ l_{13} \ (\underline{b}_3 - \underline{b}_1) \ &= \ \underline{0} \\
&\cdots \\
\underline{r}_1 \ + \ l_{1m} \ (\underline{b}_m - \underline{b}_1) \ &= \ \underline{0}
\end{aligned} \tag{16}$$

This set of $(m-1)$ equations may be rewritten in the following manner with the unknowns being $\underline{r}_1$ and the vector $\underline{l} = (l_{12}, \ l_{13}, \ ..., \ l_{1m})^T$:

$$\mathbf{A} \, \underline{r}_1 \ + \ \mathbf{B} \, \underline{l} \ = \ \underline{0} \tag{17}$$

$\mathbf{A}$ is a $(3(m-1) \times 3)$ matrix and $\mathbf{B}$ is a $3(m-1) \times (m-1)$ matrix such that :

$$\mathbf{A} \ = \ \begin{bmatrix} \mathbf{I}_{3\times3} \\ \mathbf{I}_{3\times3} \\ \cdots \\ \mathbf{I}_{3\times3} \end{bmatrix} \qquad \mathbf{B} \ = \ \begin{bmatrix} \underline{b}_2 - \underline{b}_1 & \underline{0}_3 & \underline{0}_3 & \cdots & \underline{0}_3 \\ \underline{0}_3 & \underline{b}_3 - \underline{b}_1 & \underline{0}_3 & \cdots & \underline{0}_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \underline{0}_3 & \underline{0}_3 & \cdots & \underline{0}_3 & \underline{b}_m - \underline{b}_1 \end{bmatrix} \tag{18}$$

and $\mathbf{I}_{3\times3}$ is the (3×3) identity matrix. The linear homogeneous system in (17) can be solved by means of a standard minimization with the quadratic constraint $\underline{r}_1^T \underline{r}_1 = 1$, that is with the minimization of the following criterion:

$$\mathcal{C} \ = \ \| \, \mathbf{A} \, \underline{r}_1 \ + \ \mathbf{B} \, \underline{l} \, \|^2 \ + \ \mu \, (1 - \underline{r}_1^T \underline{r}_1) \tag{19}$$

the solution of that minimization ($\frac{\partial \mathcal{C}}{\partial \underline{r}_1} = 0$ and $\frac{\partial \mathcal{C}}{\partial \underline{l}} = 0$) for $\underline{r}_1$ is given by the eigenvector corresponding to the smallest positive eigenvalue $\mu^\star$ of:

$$\mathbf{A}^T \, [\, \mathbf{I}_{3(m-1)\times3(m-1)} \ - \ \mathbf{B} \, (\mathbf{B}^T \, \mathbf{B})^{-1} \, \mathbf{B}^T \,] \, \mathbf{A} \, \underline{r}_1 \ = \ \mu \ \underline{r}_1 \tag{20}$$

Once $\underline{r}_1$ is computed, $d_\perp$ may be estimated with the previous relation (15). Since this distance does not depend on the λ_i's, the rigidity constraint may be partially relaxed. In fact, collinear projected spots are free to move all along the line supporting them. Furthermore, in our approach, the cross-sectional profile of the web is assumed to be straigth all along the web width. However, we notice that since a 1D projective basis can be defined with only three collinear points, it's not necessary to insure the straigthness all along the web width (the system of $3(m-1)$ linear equations in (17) must be of rank $(3 + (m-1))$, that is m must be greater than or equal to 3).
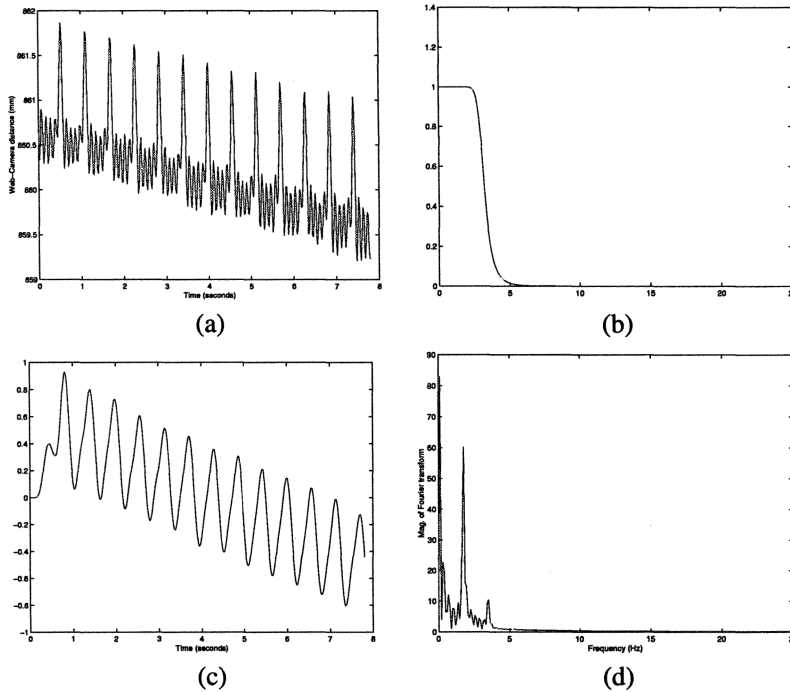
(a) (b)

(c) (d)

Figure 6: - (a) : results corresponding to the distance computation (in mm) between the camera reference frame and the centroid of the first set of collinear laser dots is reported for only 8 seconds (50 images per second, 400 images) - (b) : frequency response of the Butterworth lowpass filter ($f_c = 3$ Hz). - (c) : distance signal after filtering (*forced vibration*). The exciter is a quasi sinusoïdal function with the mean decreasing as the roll radius is decreasing. - (d) : spectrum of signal in c. The main peak is located at $f = 1.95$ Hz.

EXPERIMENTAL RESULTS

Experimental investigations have been performed on a winding plant which consists of a winder, a traction motor and a unwinder (see figure 11). For experiments, the web velocity is setting to $V = 100$ m/min and its nominal tension value is $T = 10$ N. The image acquisition system consists of an interlaced JAI CV-M50 (512 × 572) gray-scale CCIR camera connected to an Imaging Technology PCVision frame grabber at a video rate of 25 images per second. An Opton Laser optical head with a (19 × 19) dot-matrix module connected to a laser source (with wavelength of 650 nm) is also used and placed above the web. An exciter has been realized by adding an artificial non-circularity in the rolling winder. The first theoretical natural frequency of the transverse vibration has been estimated and compared with the angular velocity of the rolling winder [4]. A sequence of 1500 images (60 seconds) of 256 grey levels has been acquired on line and de-interlaced on the fly. The resulting frames (50 frames per second with 512 × 286 pixels) are stored in the computer central memory (this is about 440 Mo of memory space).

In Figure 6, results corresponding to the distance computation between the camera reference frame and the centroid of the first set of collinear laser dots is reported for only 8 seconds (the roll is being winded since this average distance decreases).
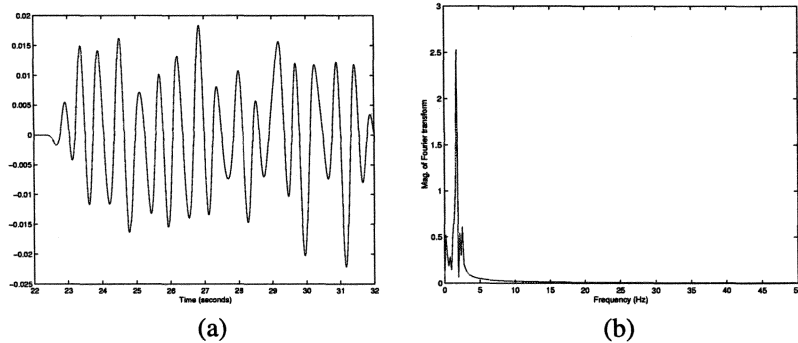
(a)   (b)

Figure 7: - (a) : web tension after filtering with a Butterworth lowpass filter (f_c = 2.5 Hz) reported for 10 seconds. - (b) : spectrum of signal in a. The main peak of this spectrum is also at 1.95 Hz. This web tension measurement by means of a piezoelectric tension sensor closed to the roll winder confirms the measurement with the vision system.

The convolution of the distance variations signal with a lowpass filter (f_c = 3 Hz) provides the forced vibration frequency (see figure 6). A value of 1.95 Hz has been found during the short sequence either with the vision system and a tension sensor (see figure 7). Other convolutions with two highpass filters (f_c = 6 Hz and f_c = 15 Hz) provide the two first modes of the free vibration of the string-behavior web (see figures 9 and 12). However, distance variations (less than 1 mm) are very small and noise is significant in the filtered signal. That's why, the resulting spectra are spread out. The first mode is precisely located in the frequency domain thanks to an elliptic passband filter (this avoids the presence of other modes as one can see in figure 10). Nevertheless, the free vibration estimated frequencies are 8.79 Hz for the first mode and about 17.5 Hz for the second mode. Other modes are out of the vision system bandwidth.
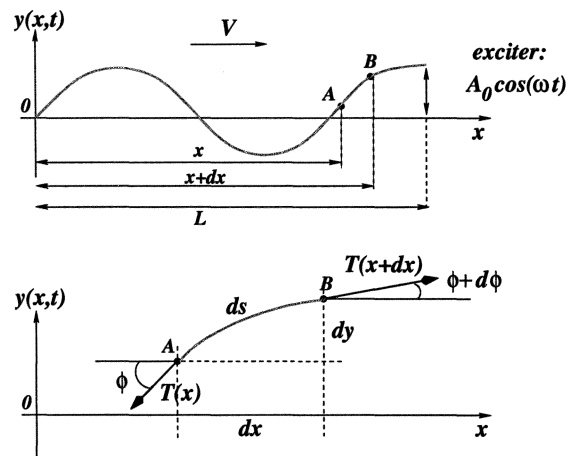

Figure 8: - The string model used for the web. One end is free whereas the other is subjected to a quasi harmonic exciter (artificially carried out with a non-circular rolling winder).

VIBRATIONS STUDY AND DISCUSSION

In this part, we discuss about the natural frequencies computed with an elastic string model for the web, and compare them with those detected with the current vision system. The transverse displacement $y(x, t)$ of the string is assumed to be small (see figure 8). The uniform elastic string of length $L = 1.378$ m (between the tractor and the roll winder), of mass density $\rho_l = 0.01375$ kg/m and with a web velocity $V = 1.666$ m/s, is subjected to a longitudinal force T. One end is free whereas the other is subjected to a small transverse displacement $y(x = L, t) = A_0 \cos(\omega t)$ where ω is the pulsation of a quasi-harmonically exciter (during a short time of the images sequence, $\omega \approx cte$). Between two disturbances, the system is left to vibrate on its own, leading to free vibrations, corresponding to internal dynamics.If the string is uniform and the tension is constant, we obtain the free vibration equation [1]:

$$(V^2 - c^2) \frac{\partial^2 y}{\partial x^2} + 2V \frac{\partial^2 y}{\partial x \partial t} + \frac{\partial^2 y}{\partial t^2} = 0 \qquad (21)$$

where $c = \sqrt{T/\rho_l} \approx 27$ m/s is the wave velocity and the nth natural frequency is:

$$f_n = n \frac{c}{2L} \left(1 - \frac{V^2}{c^2}\right) \longrightarrow f_1 \approx 9.18 \, \text{Hz} \quad \text{and} \quad f_2 \approx 18.36 \, \text{Hz} \qquad (22)$$
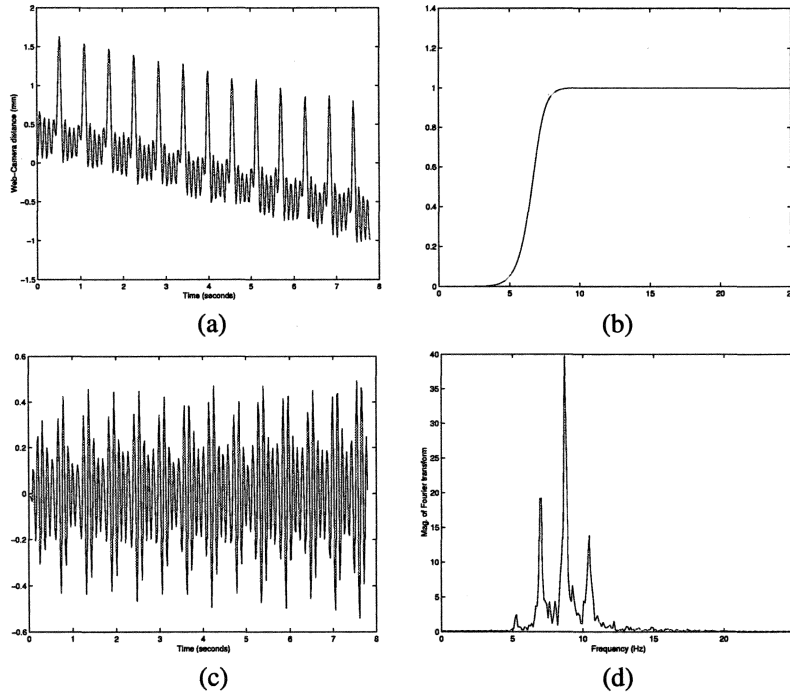

(a)

(b)

(c)

(d)

Figure 9: - (a) : results corresponding to the distance variations (in mm) between the camera reference frame and the centroid of the first set of collinear laser dots is reported for only 8 seconds. - (b) : frequency response of the Butterworth highpass filter ($f_c = 6$ Hz) . - (c) : distance variations signal after filtering (*free vibration, first mode*). - (d) : spectrum of signal in c. The main peak of this spread spectrum is at 8.79 Hz.

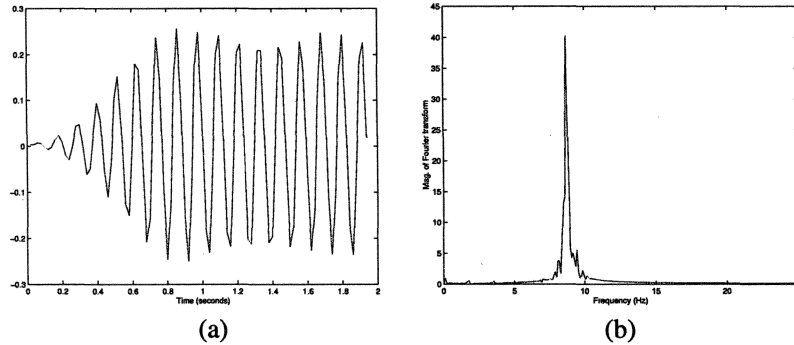(a)   (b)

Figure 10: - (a) : distance variations after filtering with an elliptic passband filter [8,10] Hz (*free vibration, again first mode*, reported for 2 seconds). This mode is well-located in the frequency domain. - (b) : the main peak of this narrow spectrum is at 8.79 Hz.

The experimental results show a good agreement with theoretical results predicted with the string model (the estimated value of the first mode is 8.79 Hz instead of 9.18 Hz). However, due to very fast and small variations of vibrations magnitude, the other modes could not be well-estimated. This is mainly due to the limited bandwidth of the current vision system (50 frames per second). Furthermore, since the tension applied in one direction influences the other direction (in the warp direction) and thereby modifies the web behavior (it's of course not a perfectly flexible string), a future work must include this biaxial tension. This also means that some pre-defined assumptions like the straigthness of the web in the width direction could be relaxed. We also think that some specific tools in the field of image segmentation and dedicated to vibrations must be further developed, by following some strategies like those carried out for video tracking applications.
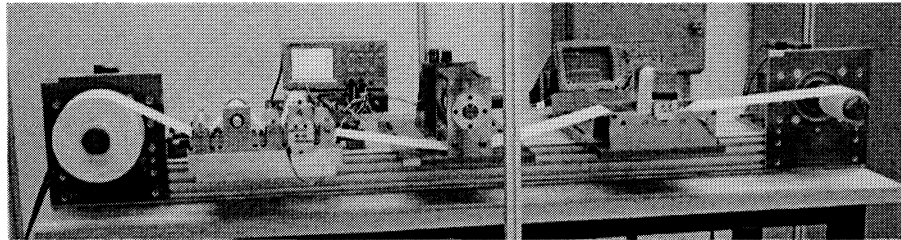

Figure 11: - The experimental winding equipment. For measurements, the motor of the rolling winder (on the right) was moved to the right in order to provide a long span of length $L = 1.378$ m (not seen on that figure, see also figure 1).

CONCLUSION

This paper presents a new measurement technique for the detection of transverse web vibrations, in motion, by means of a vision system which consists on a digital camera and a laser dots pattern device. We describe several steps for achieving a good location of laser spots, first, in the image and finally, in the 3D space. The assumptions underlying this work

have been clearly explained and some of them have been partially relaxed. We have shown that forced vibrations and the first mode of free vibrations were well detected. A range for the second mode of free vibrations has been also estimated despite the very small magnitude of vibrations in this case (less than 0.5 mm). The main advantage of this technique is to be an alternative method for detecting web vibrations without contact. In counterpart and with respect to other sensors used in the field of winding systems, vision systems generally have a very limited bandwidth and require cumbersome computations for extracting the video informations. Nevertheless, the good agreement of the results with the theory and the precise location of laser spots (that is, a sampling area of the web surface) during the winding process should be lead in the future to a promising measurement technique.
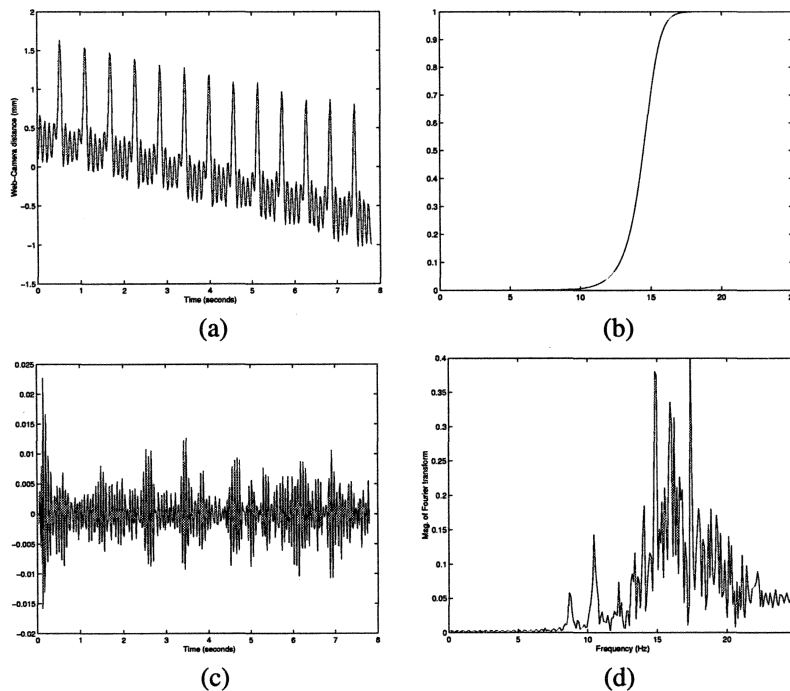
ACKNOWLEDGMENTS

(a)

(b)

(c)

(d)

Figure 12: - (a) : results corresponding to the distance variations (in mm) between the camera reference frame and the centroid of the first set of collinear laser dots is reported for only 8 seconds. - (b) : frequency response of the Butterworth highpass filter ($f_c = 15$ Hz) - (c) : distance variations signal after filtering (*free vibration, second mode*). - (d) : spectrum of signal in c. It's a bell-shaped spectrum. At this scale (distance variations are less than 0.5 mm), noise is significant. However, the main peak is located at about 17.5 Hz.

References

[1] Aubry, E., "Aspects dynamiques et vibrations des systèmes", Mémoire d'Habilitation à diriger des recherches, December 1997, Mulhouse, France.

[2] Castleman, K., "Digital Image Processing", Prentice Hall, ISBN: 0-13-211467-4, 1996.

[3] Doignon, C., Abba, G., "A practical multi-plane method for a low-cost calibration technique", Proceedings of the fifth European Control Conference, Karlsruhe, Germany, Sept. 1999.

[4] Doignon, C., Knittel, D., "Non-circularity detection of a rolling winder by artificial vision", Proceedings of the fifth International Conference on Quality Control by Artificial Vision, May 2001, Le Creusot, France.

[5] Duda, R.O., Hart, P.E., "Pattern Classification and Scene Analysis", Wiley interscience publication, ISBN 0-471-22361-1, New-York, 1973.

[6] Gueldenberg, B., Welp, E.G., "Quantitative analysis of nip-induced tension by use of digital image processing", Proceedings of the International Conference on Web Handling, IWEB5, Stillwater, Oklahoma, USA, 1999.

[7] Klette, R., Karstein, S., Koschan, A., "Computer Vision : Three-Dimensional Data from Images", Springer-Verlag, ISBN 981-3083-71-9, Singapour, 1998.

[8] Koc, H., Knittel, D., De Mathelin, M., and Abba, G., "Modeling and robust control of winding for elastic webs", IEEE Transactions on Control System Technology, Vol. 10, $n^o 2$, pp. 197-208, April 2002.

[9] Lakshmikumaran, A.V., Wickert, J.A., "Edge buckling of imperfectly guided webs", Transactions on ASME Journal of Vibration and Acoustics, Vol.120 April, pp. 346-352, 1998.

[10] Longuet-Higgins, H.C., "A computer algorithm for reconstructing a scene from two projections", Nature, vol. 293, pp. 133-135, September 1981.

[11] Reid, K.N., Lin, K.C., "Control of longitudinal tension in multi-span web transport systems during start-up", Proceedings of the International Conference on Web Handling, p. 77-95, IWEB2, Stillwater, Oklahoma, USA, 1993.

[12] Haralick, R.M., Shapiro, L.G., "Computer and Robot Vision", Addison-Wesley Publishing, vol. n° 2, ISBN 0-201-10877-1, 1992.

[13] Shrikhande, N., Stockman, G., "Surface Orientation from a Projected Grid", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.11, n° 6, June 1989.

[14] Wang, Y.F., Mitiche, A., Aggarwal, J.K., "Computation of Surface Orientation and Structure of Objects using Grid Coding", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 9, n° 1, January 1987.

[15] Wolfermann, W., "Compensation of disturbances in the web force caused by a non-circular running winder", Proceedings of the International Conference on Web Handling, IWEB5, Stillwater, Oklahoma, USA, 1999.