

DEVELOPMENT OF A NEW LASER-BASED, OPTICAL INDOOR POSITIONING SYSTEM

S. Tilch, R. Mautz

Institute of Geodesy and Photogrammetry, ETH Zurich, 8093 Zurich, Switzerland (sebastian.tilch,
rainer.mautz)@geod.baug.ethz.ch

Commission V, WG V/1

KEY WORDS: Relative Camera Orientation, Point Algorithm, Inverse Camera Principle, Indoor Positioning System

ABSTRACT:

This paper presents a novel optical indoor positioning system that is currently under development at the Institute of Geodesy and Photogrammetry at the ETH Zurich. The aim of the project is to develop an automatic, inexpensive method capable of providing continuous sub-mm positions in real-time from a mobile device in all indoor environments. The system architecture does not require cost-intensive high-precision mechanics or sophisticated set-ups. The main purpose of the system called CLIPS (Camera and Laser based Indoor Positioning System) is the precise pose determination of a mobile camera. The difference to the photogrammetric relative camera orientation is that it is based on the principle of an inverse camera, where one camera is substituted for a device that emits laser-beams from a virtual central point. The laser beams project bright spots on any surface in an indoor environment. These laser spots are seen by a camera for the determination of its relative orientation with respect to the laser-device. First experiments have shown that the relative orientation of the camera could be correctly determined in all cases and that the new system has the potential to achieve mm-level accuracy or better. However, the overall system performance has been limited so far due to an imprecise determination of the system-scale.

1. INTRODUCTION

A diverse range of indoor positioning technologies is already on the market or currently under development. Techniques that have the potential to achieve cm-level accuracy or better include the use of Ultra-Wideband (UWB), WLAN, Ultrasound, RFID, high sensitive GNSS, pseudolites and many more. However, one single system cannot satisfy all user requirements due to poor signal propagation or restrictions in the line of sight. In particular, some systems require sophisticated set-ups or fix installations and therefore do not fulfil the users' requirements of flexibility and mobility. On the other hand, systems that operate without additional infrastructure, i.e. signal strength based methods, cannot achieve sub-decimetre accuracy and do not reach the level of reliability that the user needs. Our novel optical system CLIPS (Camera and Laser based Indoor Positioning System) proposed for the first time in Mautz (2009) aims to overcome these drawbacks.

2. ARCHITECTURE OF CLIPS

The central idea of the new indoor positioning system is based on the fundamentals of stereo photogrammetry, where the position and the rotation of a camera relative to another camera are derived. Instead of using two cameras, the principle of an inverse camera is used by replacing one camera with a device that we call "laser-hedgehog". This device projects well-distributed laser spots as flexible reference points on the ceiling, walls and furnishings in any indoor environment. The projecting light source consists of several focused laser-beams that originate from a static, well-defined central point. The 3D directions of the laser-beams are also precisely known through a one-time high precision calibration. By projecting the laser-beams on a virtual (i.e. mathematical) plane, we are able to

simulate the image of a virtual camera. The main functions of the laser-hedgehog can be summarised as:

1. Projection of flexible reference points on any surface.
2. Simulation of a second camera.
3. Allowing for the use of a computational cheap point detection algorithm.

The advantages of that approach are threefold. Firstly, the system is not depending on an existing field of reference points. The laser-hedgehog simply creates its own reference field. As a consequence, CLIPS has a high degree of mobility that is quickly and immediately applicable in any indoor environment. Secondly, the computational costs can be reduced since there is only one camera image that has to be processed. Through a one-time calibration of the laser-hedgehog, the orientation of each laser beam is already known within 0.05° in its horizontal and vertical angle. As all laser beams have a common origin, we can use stereographic projection to map the directions of the lasers on a virtual image. Thus, we spare the steps of point detection and identification for the virtual picture. Thirdly, the detection of corresponding point pairs is facilitated with the laser spots due to their distinct colour, shape and brightness compared to feature detection algorithms that rely on natural image features.

During the measurement phase the digital camera observes the reference field, which is in our case projected by the laser-hedgehog. Having carried out the point detection and identification for the individual laser beams in the camera image, the relative orientation (detailed in Chapter 3) can be computed by introducing the coplanarity constraints of epipolar geometry.

Another problem is the introduction of a system-scale which cannot be determined by relative orientation. Out of several

options for the introduction of the system scale, the simplest option was chosen by directly measuring the distance between the laser hedgehog and the camera using another instrument. These distance measurements were carried out only for the first couple camera positions. Then, the 3D vectors between the laser hedgehog and the camera positions were computed. In addition, the spatial coordinates of the laser points were determined by intersection. Once the 3D positions of the laser spots are known, the relative orientation parameters for further camera positions can be determined by simple spatial resection.

3. RELATIVE CAMERA ORIENTATION

When determining the geometric relationship between the laser hedgehog and the CLIPS camera the presence of known 3D reference points as well as the provision of prior approximate values for the relative camera orientation cannot be assumed. This poses a challenging problem for the determination of the unknown orientation parameters. Therefore, the requirement for an algorithm is to be capable to correctly, uniquely and robustly determine the relative orientation even in absence of an approximate guess of the initial camera pose. In addition, the to-be-used algorithm should perform well with a minimum number of pairs of corresponding points and should not have any restrictions to the geometry of the projected laser points. These requirements are the criteria for assessing the relative orientation algorithm that is best qualified for CLIPS.

| Non-iterative Methods | | | |
|--------------------------------------------|---------------------|---------------------|---------|
| | 5-Point (Stewénius) | 6-Point (Stewénius) | 8-Point |
| geometry used | epipolar geometry | | |
| number of point-correspondences | 5 | 6 | 8 |
| inner camera orientation needs to be known | yes | no | yes |
| control points required | no | no | no |
| least squares adjustment | no | no | no |
| initial approximate guess required | no | no | no |
| number of solutions | 1-10 | 1-15 | 1 |

Table 1. Non-iterative methods for the determination of the relative camera orientation

The 5-Point and the 6-Point Algorithms have two important advantages: a) they do not require approximate values for the relative orientation and b) they perform well even if the point-correspondences are entirely on a critical surface such as a plane. As a drawback both algorithms do not deliver unique solutions, but instead a couple of possible solutions where the correct solution needs to be identified. The 8-Point Algorithm is a linear method that does not require an approximate solution and yields a unique solution, but fails in case that all point-correspondences are located on the same plane.

Iterative methods require the provision of initial approximate solutions for the unknown relative orientation parameters that get refined by an iterative least-squares adjustment. If these series of iteratively refined solutions converges, a unique solution will be the outcome.

| Iterative Methods | | |
|--------------------------------------------|--------------------------------------------|-------------------|
| | exploitation of the coplanarity constraint | spatial resection |
| geometry used | epipolar geometry | collinearity |
| number of point-correspondences | 5 | 3 |
| inner camera orientation needs to be known | yes | yes |
| control points required | no | yes |
| least squares adjustment | yes | yes |
| initial approximate guess required | yes | yes |
| number of solutions | 1 | 1 |

Table 2. Iterative methods for the determination of the relative camera orientation

The comparison between the requirements stated above and the properties given in Table 1 and 2 shows, that none of the algorithms can fulfil all criteria. Since the inner orientation, i.e. the intrinsic camera parameters, can assumed to be known for the CLIPS camera, the 6-Point Algorithm is not adequate. Also the 8-point algorithm is not a feasible algorithm for CLIPS, because it cannot be applied if all points are located on a plane, which is likely to be the case if all laser-points are projected to the ceiling. For this reason, the chosen method for the relative orientation consists of a catenation of the three remaining algorithms: the 5-Point Algorithm provides the initial approximations that are optimised by an iterative least-squares minimisation based on the coplanarity and collinearity constraints. As the 5-Point Algorithm provides up to 10 solutions for every camera position, the algorithm has been embedded into a RANSAC algorithm. As a result, the correct essential matrix is identified. Then, the correct solution is decomposed into a translational vector \underline{b} and a rotational matrix R and finally adjusted by a least-squares estimation of all relative orientation parameters. This approach is illustrated in Figure 1.

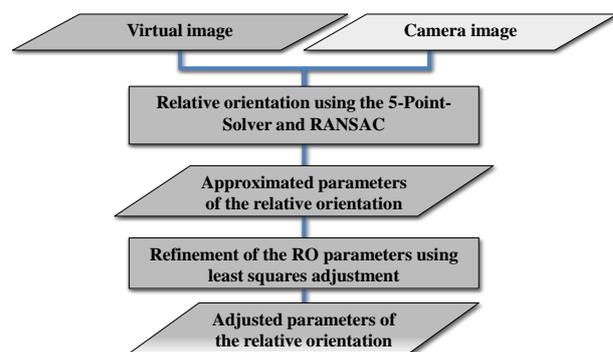


Figure 1. Method for the estimation of the relative camera orientation

3.1.1 5-Point Algorithm of Stewénius

The 5-Point Algorithm that we use is based on Stewénius (2005). The use of this algorithm requires a previous inner camera orientation that can be obtained by a one-time camera calibration. Note that the 5-Point Algorithm developed by Stewénius does not estimate the camera pose that is normally expressed in form of the base vector \underline{b} and the orientation

angles (roll, pitch und yaw). The algorithm yields the essential matrix

$$E = R'^T B R'' \quad (1)$$

that generally describes the relative orientation (R' , R'' , B) of two cameras with respect to each other, where R' is the rotation matrix of the inverse camera and R'' the rotation matrix of the real camera. The base matrix

$$B = \begin{pmatrix} 0 & -b_z & b_y \\ b_z & 0 & -b_x \\ -b_y & b_x & 0 \end{pmatrix} \quad (2)$$

is skew-symmetric and contains the components of the base vector $\underline{b} = (b_x, b_y, b_z)^T$ between the projection centres of the two cameras. Under the assumption that the inverse camera (in our case the laser-hedgehog) has already been set-up in the target coordinate system, the rotation matrix R' is equal to the identity matrix I . Thus the equation for the essential matrix

$$E = B R'' \quad (3)$$

is simply the product of the base matrix B and the rotation matrix R'' of the digital camera. In order to compute the essential matrix E the coplanarity constraint

$$\underline{x}''^T E \underline{x}' = 0 \quad (4)$$

is applied. The coplanarity condition exploits the fact that the projection centre of the camera O' , the projection centre of the inverse camera O'' and each laser point are on the same plane, called the epipolar plane. The coordinates \underline{x}' and \underline{x}'' denote the image coordinates of the inverse camera and the digital camera. Alternatively, the coplanarity constraint can be written as a scalar triple product

$$\underline{r}'^T (\underline{b} \times \underline{r}'') = \underline{r}'^T B \underline{r}'' = 0 \quad (5)$$

The image projection vectors in the object coordinate system

$$\underline{r}' = R' \underline{x}' = I \underline{x}' = \underline{x}' \quad \text{and} \quad \underline{r}'' = R'' \underline{x}'' \quad (6)$$

of a world point L correspond to the translated and rotated image vectors \underline{x}' and \underline{x}'' that appeared in (4).

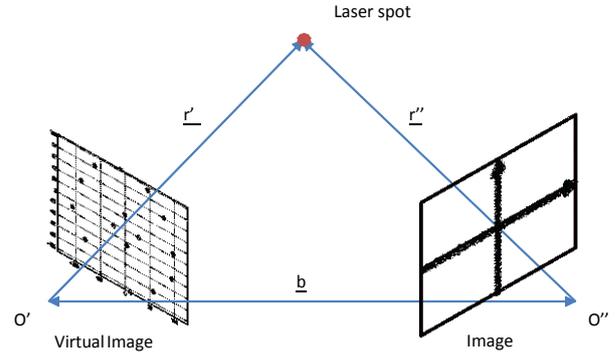


Figure 2. Epipolar geometry

In order to compute the essential matrix E , the coplanarity constraint (4) is modified to

$$A \underline{e} = 0 \quad (7)$$

where

$$A = \begin{pmatrix} x_1' x_1'' & \dots & y_1' x_1'' & \dots & z_1' z_1'' \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n' x_n'' & \dots & y_n' x_n'' & \dots & z_n' z_n'' \end{pmatrix} \quad (8)$$

contains as many lines as there are point correspondences (Hartley et al. 2003). Vector \underline{e} contains all 9 elements of the essential matrix E . A singular value decomposition (SVD) is performed on the A -matrix by analysing

$$USV^T = A \quad (9)$$

The four column vectors of matrix V that have the smallest singular values are decomposed into the matrices E_1, E_2, E_3 and E_4 . Thus, the essential matrix E can be expressed as a linear combination in the form of

$$E = xE_1 + yE_2 + zE_3 + wE_4 \quad (10)$$

where x, y, z and w are scalars. Since these four variables can be determined apart from one unknown scale factor, w is set to 1. Using the way Stewénius (2005) has chosen to solve the equation system, the two additional constraints

$$\det(E) = 0 \quad (11)$$

$$2EE^T E - \text{trace}(EE^T)E = 0$$

are introduced. The expansion of the essential matrix E according to (10) in (11), results in a polynomial system of 10 equations of degree 3. This set of equations can be solved with

the Gröbner bases where a 10 x 10 matrix is set up. The eigenvectors of that matrix contain the solutions for the unknowns x , y and z of (10). This way, all possible essential matrices can be expressed. A detailed description of the algorithm can be found in Stewenius (2005).

The advantage of that approach is that theoretically only five corresponding pairs of object points need to be identified in order to estimate matrix E and obtain a solution for the relative orientation. A critical geometric configuration occurs if all object points are elements of a straight line. In addition, Kalantari (2009) pointed out that the 5-Point Algorithm by Stewenius is not robust for the case of sideways motion. The algorithm is useful for the CLIPS project, because it performs well even if all corresponding points are on the same plane.

3.1.2 Combination of the 5-Point Algorithm with MSAC

The drawback of the 5-Point Algorithm is that the result can consist of up to ten solutions in form of different E -matrices. In order to identify the correct matrix in real time, an MSAC (M-Estimator SAmple Consensus) Algorithm according to Torr et al. (2000) has been implemented. In order to decide whether a point pair is an outlier or inlier, the cost function

$$C = \sum_{i=1}^n \rho(w_i) \quad (12)$$

is set up, where the weight function reads

$$\rho(w_i) = \begin{cases} w_i & w_i < T \\ T & w_i \geq T. \end{cases} \quad (13)$$

In contrast to the RANSAC algorithm, not only outliers ($w_i \geq T$) contribute to the cost function C , but also the inliers ($w_i < T$). The weight of each outlier is equal to a pre-defined threshold T . Inliers also contribute to the function costs with an inconsistency value w that is determined by

$$w_i = \left| \underline{x}_i^T E \underline{x}_i'' \right| \neq 0. \quad (14)$$

An individual threshold T is determined for every essential matrix E separately.

Prior to the MSAC test, geometric impossible essential matrices E are eliminated by a plausibility check on the geometry. Obviously, if the object points that have been determined by spatial intersection between the vectors \underline{r}_i' and \underline{r}_i'' are located behind the laser-hedgehog, the solution on hand can be rejected. The essential matrix that is most likely the correct solution according to the MSAC test, will be further considered. In order to carry out the MSAC test, each essential matrix is decomposed in rotation matrix R'' and base matrix B according to Hartley et al. (2003). First, a singular value decomposition is performed on the essential matrix E by

$$E = U \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} V^T. \quad (15)$$

Introducing the matrices

$$W = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } Z = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (16)$$

the base matrix and rotation matrix can be determined by

$$B_1 = UZU^T \text{ and } B_2 = -B_1, \quad (17)$$

$$R_1'' = UWV^T \text{ and } R_2'' = UW^T V^T.$$

As mentioned in Hartley et al. (2004) four different camera poses can be derived from each essential matrix E . This ambiguity problem arises from the fact that the essential matrices are only unique up to an undefined algebraic sign that induces two solutions for the base and the rotation matrix respectively. However, with the application of the previously mentioned geometric check, the three false solutions can be identified. The geometric check is based on a spatial intersection according to Tilch (2010). Hereby, the two normalized vectors

$$\underline{n}_i' = \frac{\underline{r}_i'}{\|\underline{r}_i'\|} \text{ and } \underline{n}_i'' = \frac{\underline{r}_i''}{\|\underline{r}_i''\|} \quad (18)$$

are determined and with the help of the matrices

$$A_i' = I - \underline{n}_i' \underline{n}_i'^T \quad (19)$$

$$A_i'' = I - \underline{n}_i'' \underline{n}_i''^T$$

the design matrix

$$A_i = \begin{pmatrix} A_i' \\ A_i'' \end{pmatrix} \quad (20)$$

can be set up, where matrix I is a 3 x 3 identity matrix. The observation vectors

$$\underline{l}_i' = A_i' \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (21)$$

$$\underline{l}_i'' = A_i'' \underline{b}$$

are here written as a product of the origin of the position vectors and the matrices A_i' respectively A_i'' . The coordinate origin $(0, 0, 0)^T$ was chosen to be the initial point for the position vectors of all laser rays L_i' and \underline{b} to be the initial point for L_i'' accordingly. Both observation vectors can be assembled as a common observation vector

$$L_i = \begin{pmatrix} L_i' \\ L_i'' \end{pmatrix}. \quad (22)$$

The 3D position vector of the laser points \underline{X}_i can be determined by a least-squares adjustment

$$\underline{X}_i = (A_i^T A_i)^{-1} (A_i^T L_i) \quad (23)$$

apart from an unknown system scale factor. For the application of the geometric check, this procedure is repeated for every combination of possible base vectors \underline{b} and rotation matrices R .

3.1.3 Refinement of the Camera Pose by Least-Squares Minimisation

For the final determination of the relative orientation parameters, an iterative least-squares adjustment is carried out. Such an adjustment requires good approximate initial values for base vector \underline{b} and the rotation matrix R , which can be taken from the outcome of the MSAC Algorithm. Note that such a procedure requires the explicit determination of the rotation angles

$$\omega = \arctan\left(\frac{-r_{23}}{r_{33}}\right), \varphi = \arctan\left(\frac{r_{13}}{\sqrt{r_{11}^2 + r_{12}^2}}\right), \quad (24)$$

$$\kappa = \arctan\left(\frac{-r_{12}}{r_{11}}\right)$$

where the components of the rotation matrix read

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}. \quad (25)$$

The observation equations are derived from the coplanarity constraint (5) and conveniently rewritten as determinants

$$\Delta = \begin{vmatrix} 1 & x' & r_x'' \\ b_y^* & y' & r_y'' \\ b_z^* & z' & r_z'' \end{vmatrix} = 0 \quad (26)$$

analogous to Luhmann (2003). The entries of the first column in (26) are obtained by a division of the original constraints with the base vector component b_x and the substitution

$$b_x^* = 1, b_y^* = \frac{b_y}{b_x} \text{ and } b_z^* = \frac{b_z}{b_x}. \quad (27)$$

The position vector

$$\underline{r}'' = (r_x'' \ r_y'' \ r_z'')^T \quad (28)$$

denotes the direction vectors \underline{x}'' after the camera coordinate system has been rotated into the direction of reference coordinate system analogous to (6) using the rotation matrix R'' . From the observation equations (26) we obtain one nonlinear constraint for each point correspondence. In order to apply the iterative Gauss-Newton optimisation these constraints need to be linearised. Therefore the nonlinear observation equations (26) are transformed into the linear model

$$v_\Delta = \frac{\partial \Delta}{\partial b_y} db_y + \frac{\partial \Delta}{\partial b_z} db_z + \frac{\partial \Delta}{\partial \omega} d\omega + \frac{\partial \Delta}{\partial \varphi} d\varphi + \frac{\partial \Delta}{\partial \kappa} d\kappa + \Delta^0, \quad (29)$$

where each addend denotes the partial derivative of (26) in respect to one of the unknown orientation parameters $b_y^*, b_z^*, \omega, \varphi$ and κ . For each available point correspondence one observational equation (29) is obtained. Therefore at least five point correspondences are necessary in order to solve for all five unknown parameters. The coefficients of (29) are the entries of a design matrix A that is used for the Gauss-Newton least-squares adjustment. Accordingly, the normal equations read

$$\underline{dx} = N^{-1} n, \quad (30)$$

where

$$N = A^T A \quad (31)$$

$$n = A^T \Delta.$$

The vector \underline{dx} in (30) contains an additive correction for the approximate values of the orientation unknowns that are taken from the outcome of the previous iteration step. If the entries in \underline{dx} do not contribute to a significant improvement anymore, the iteration is stopped. As a result of this refinement we obtain the adjusted unknown relative orientation parameters, i.e. the 3D translation vector and the rotation angles between the digital camera and the laser-hedgehog. Along with a probe attached to the camera, the translation vector can be used for coordinate measurements on industrial objects.

4. ASSESSMENT OF THE PROTOTYPE

The accuracy assessment of the current system can only be regarded as provisional because the projection centre of the

camera used was neither made available by the manufacturer nor straightforwardly determinable. As a workaround, the accuracy of the 3D-coordinates of the laser points has been assessed. Based on the recording of the laser points from various camera locations and determination of the camera orientation for each location as described in Chapter 3, the 3D-coordinates of the laser point centres could be determined by spatial intersection. In order to obtain an independent and redundant set of laser point coordinates, a totalstation survey has been carried out. The comparison was carried out by assessing the residuals of an affine transformation between both 3D-point clouds. Hereby, the standard deviation σ_0 and the scale factor m of the affine transformation have been considered as the main evaluation criteria. The 1-sigma standard deviation was $\sigma_0 = 0.6$ mm and the scale factor $m = 0.9981$ that would cause deviations of 2 mm/m. Our explanation for the relative large difference from the expected scale factor of $m = 1$ is that the exact location of the camera projection centre was unknown. In order to assess the precision (i.e. the repeatability under unchanged conditions) of our measurement system we have taken images at five different camera positions, each with 20 camera shots. The average 1- σ standard deviation of a single measurement was $\sigma_x = 0.08$ mm, $\sigma_y = 0.08$ mm and $\sigma_z = 0.18$ mm. We ascribe the larger variance in the height component to the geometry of the laser points that were all projected on the ceiling of our laboratory.

5. CONCLUSIONS AND OUTLOOK

The relative orientation of the CLIPS camera could be successfully achieved with an implementation of the 5-Point Algorithm of Stewénius (2005). Through a subsequent least-squares adjustment the pose estimation could be further refined. First tests have proved that the pose estimation of the camera is reproducible. Due to missing knowledge of the camera projection centre, a throughout system assessment was not possible. Nevertheless, the tests show that CLIPS has the potential to achieve a positioning accuracy in the magnitude of millimetres or better.

If the number of laser points could be increased, the 5-Point Algorithm would perform more stable in respect to measurement noise and outliers. Along with the increase in the number of laser points the next steps in the system development include the improvement of the laser point detection algorithm and an enhanced identification method. Furthermore, a proper and practicable solution for the determination of the system scale must be found.

6. REFERENCES

- Hartley R., Zisserman A. (2003): *Multiple view geometry in computer vision*, 2nd Edition. Cambridge University Press, Cambridge.
- Kalantari, M., Jung, F., Guedon, J. and Paparoditis, N. (2009): The Five Points Pose Problem: A New and Accurate Solution Adapted to Any Geometric Configuration, *Lecture Notes In Computer Science* Vol. 5414, pp. 215-226 (2009).
- Luhmann, T. (2003): *Nahbereichsphotogrammetrie – Grundlagen, Methoden und Anwendungen*, 2nd Edition. Herbert Wichmann Verlag, Hüthig GmbH & Co. KG, Heidelberg.
- Mautz, R. (2009): Camera and Laser Indoor Positioning System (CLIPS). Proposal to the Swiss National Science Foundation, (unpublished).
- Stewénius, H. (2005): Gröbner Basis Methods for Minimal Problems in Computer Vision, Ph.D. thesis, Centre for Mathematical Sciences, Lund Institute of Technology.
- Tilch, S. (2010): Entwicklung eines optischen Innenraum-Positionierungssystems, Master Thesis, Institute of Geodesy and Photogrammetry, ETH Zurich, 2010.
- Torr P., Zisserman A. (2000): MLESAC: A new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding* 78, pp. 138-156 (2000).