Region-based motion-compensated iterative reconstruction technique for dynamic computed tomography

REGION-BASED MOTION-COMPENSATED ITERATIVE RECONSTRUCTION TECHNIQUE FOR DYNAMIC COMPUTED TOMOGRAPHY

Anh-Tuan Nguyen, Jens Renders, Jan Sijbers, and Jan De Beenhouwer

imec-Vision Lab, Department of Physics, and μNEURO Research Centre of Excellence
University of Antwerp, Universiteitsplein 1, Wilrijk 2610, Belgium

ABSTRACT

Current state-of-the-art motion-based dynamic computed tomography reconstruction techniques estimate the deformation by considering motion models in the entire object volume although occasionally the proper change is local. In this article, we address this issue by introducing the region-based Motion-compensated Iterative Reconstruction Technique (rMIRT). It aims to accurately reconstruct the object being locally deformed during the scan, while identifying the deformed regions consistently with the motion models. Moreover, the motion parameters that correspond to the deformation in those areas are also estimated. In order to achieve these goals, we consider a mathematical optimization problem whose objective function depends on the reconstruction, the deformed regions and the motion parameters. While dynamic CT methods that specifically locate the regional changes in the object volume are not widely found in the literature, our method exploits the analytical derivative towards these changing regions, which allows for efficient reconstruction using gradient-based optimizers.

***Index Terms*—** Dynamic computed tomography, motion parameter estimation, region estimation.

1. INTRODUCTION

Dynamic computed tomography is a major part of CT imaging that studies the structure of dynamic objects in CT scans. Typical examples of state-of-the-art dynamic CT reconstruction techniques can be [1, 2] that jointly reconstruct the density volume and estimate the deformation vector fields (DVFs) between successive time frames, or [3, 4, 5] that estimate the motion parameters that correspond to specific motion models. A disadvantage of all those methods is they consider the motion in the entire object volumes, while in real applications (e.g., lung tissue [6]) only local regions are deformed. This concern was recently mentioned in [3]. Although several reconstruction methods were designed to estimate these local deformed regions (e.g., [7, 8]), those methods do not consider affine motions in those local areas.

In this paper, we address that issue by considering a dynamic CT model for which, in the object volume, there are local regions deformed by affine motion models, while the complementary regions that remain static during the entire acquisition scan. We then propose an iterative method that aims not only to accurately reconstruct the scanned object that contains these locally deformed regions, but also to identify them. Furthermore, the motion parameters corresponding to the deformation are estimated simultaneously with the reconstruction and region estimation. The contributions are summarized as follows:

- Formulation of the class of dynamic CT problems that consider affine motions, which model the deformation in local areas characterized by binary masks, assuming the change in the entire object volume is continuous.
- Gradient method that aims to minimize an objective function that depends on the reconstruction, the motion parameters and the deformed regions, whose partial derivatives towards all of them are formulated analytically.
- The biconvexity of the objective function towards the reconstruction and the locally deformed regions that supports the convergence of the iterative schemes in the proposed gradient method.

2. PROPOSED METHOD

A dynamic CT image can be represented as a sequence of n images $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, ..., $\boldsymbol{x}_n$, each representing the object at a given point in time. The acquisition can be observed as a collection of finite subscans, where the object is assumed to be static during each subscan. Here, a subscan refers to one or more consecutively acquired projections. This procedure can be mathematically modelled as n systems of linear equations:

$$\boldsymbol{W}_i \boldsymbol{x}_i = \boldsymbol{b}_i, \text{ for } i = 1, ..., n, \tag{1}$$

where $\boldsymbol{W}_i$ and $\boldsymbol{b}_i$ are the projection operator and the projection data corresponding to the i^{th} subscan, respectively. These may be interpreted as a single system of the forward

model:

$$\begin{bmatrix} \boldsymbol{W}_1 & 0 & 0 & 0 \\ 0 & \boldsymbol{W}_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \boldsymbol{W}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_n \end{bmatrix}. \tag{2}$$

Let $\boldsymbol{\alpha}_i \in \{0,1\}^N$ be a binary mask, which encodes the local region of the unknown original image $\boldsymbol{x} \in [0,1]^N$ that appears deformed in the image $\boldsymbol{x}_i$. Assume the local deformation can be modelled by an affine motion model M that depends on the motion parameter $\boldsymbol{p}_i \in \mathbb{R}^M$ and assume the deformation in the entire object volume is continuous, the deformed object in the i^{th} subscan then can be modelled as follows:

$$\boldsymbol{x}_i = \overline{\boldsymbol{\alpha}_i} \circ \boldsymbol{x} + M(\boldsymbol{p}_i)(\boldsymbol{\alpha}_i \circ \boldsymbol{x}), \tag{3}$$

where $\overline{\boldsymbol{\alpha}_i} := \boldsymbol{1} - \boldsymbol{\alpha}_i$ and $\circ$ is the commutative Hadamard product. In this model, the static part $\overline{\boldsymbol{\alpha}_i} \circ \boldsymbol{x}$ of $\boldsymbol{x}$ remains conserved in the deformed object $\boldsymbol{x}_i$, while the dynamic part $\boldsymbol{\alpha}_i \circ \boldsymbol{x}$ appears distorted under the motion model M. By substituting the equation (3) to (2) for all n, the forward model of the entire projection data then may be interpreted as the following single system:

$$\begin{bmatrix} \boldsymbol{W}_1 & 0 & 0 & 0 \\ 0 & \boldsymbol{W}_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \boldsymbol{W}_n \end{bmatrix} \begin{bmatrix} \overline{\boldsymbol{\alpha}_1} \circ \boldsymbol{x} + M(\boldsymbol{p_1})(\boldsymbol{\alpha_1} \circ \boldsymbol{x}) \\ \overline{\boldsymbol{\alpha}_2} \circ \boldsymbol{x} + M(\boldsymbol{p_2})(\boldsymbol{\alpha_2} \circ \boldsymbol{x}) \\ \vdots \\ \overline{\boldsymbol{\alpha}_n} \circ \boldsymbol{x} + M(\boldsymbol{p_n})(\boldsymbol{\alpha_n} \circ \boldsymbol{x}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_n \end{bmatrix} \tag{4}$$

This can be concisely rewritten as a single system:

$$\boldsymbol{W}\{\overline{\boldsymbol{\alpha}}[\circ]\boldsymbol{x} + \boldsymbol{M}(\boldsymbol{p})(\boldsymbol{\alpha}[\circ]\boldsymbol{x})\} = \boldsymbol{b}. \tag{5}$$

where

$$\boldsymbol{\alpha} = \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \vdots \\ \boldsymbol{\alpha}_n \end{bmatrix}, \boldsymbol{p} = \begin{bmatrix} \boldsymbol{p}_1 \\ \boldsymbol{p}_2 \\ \vdots \\ \boldsymbol{p}_n \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_n \end{bmatrix}, \tag{6}$$

$[\circ]$ is the modified version of the penetrating face product [9] between the two column vectors $\boldsymbol{\alpha} \in \{0,1\}^{nN}$ and $\boldsymbol{x} \in [0,1]^N$ defined by

$$\boldsymbol{\alpha}[\circ]\boldsymbol{x} = \left[[\boldsymbol{\alpha}_1 \circ \boldsymbol{x}]^T, [\boldsymbol{\alpha}_2 \circ \boldsymbol{x}]^T, \ldots, [\boldsymbol{\alpha}_n \circ \boldsymbol{x}]^T\right]^T, \tag{7}$$

and

$$\boldsymbol{M}(\boldsymbol{p}) = \begin{bmatrix} M(\boldsymbol{p}_1) & 0 & 0 & 0 \\ 0 & M(\boldsymbol{p}_2) & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & M(\boldsymbol{p}_n) \end{bmatrix}. \tag{8}$$

In order to solve the equation (5), let us consider the following constrained optimization problem as a modified and extended version of [4, 5, 10]:

$$[\boldsymbol{x}^*, \boldsymbol{\alpha}^*, \boldsymbol{p}^*] = \underset{\boldsymbol{x} \in [0,1]^N, \boldsymbol{\alpha} \in \{0,1\}^{nN}, \boldsymbol{p} \in \mathbb{R}^{nM}}{\arg\min} f(\boldsymbol{x}, \boldsymbol{\alpha}, \boldsymbol{p}), \tag{9}$$

where

$$f(\boldsymbol{x}, \boldsymbol{\alpha}, \boldsymbol{p}) = \frac{1}{2} \left\| \boldsymbol{W}\{\overline{\boldsymbol{\alpha}}[\circ]\boldsymbol{x} + \boldsymbol{M}(\boldsymbol{p})(\boldsymbol{\alpha}[\circ]\boldsymbol{x})\} - \boldsymbol{b} \right\|_2^2. \tag{10}$$

The problem (9) can be solved by the iterative schemes presented in the Algorithm 1 with the intermediate estimated value of $\boldsymbol{\alpha}$ is projected onto the non-convex set $\mathcal{S} := \{0,1\}^{nN}$ by the following projector to obtain the intermediate deformed regions, with τ is a suitable threshold:

$$\text{Proj}_{\mathcal{S},\tau}(\boldsymbol{\alpha})[i] := \begin{cases} 0 & \text{if } \boldsymbol{\alpha}[i] < \tau \\ 1 & \text{otherwise} \end{cases}, \text{for } i = \overline{1, nN}. \tag{11}$$

Algorithm 1: rMIRT

Input: Projection $\boldsymbol{b}$, projector $\boldsymbol{W}$, motion model $\boldsymbol{M}$, $\boldsymbol{p}^0 \equiv$ motion parameters in the static case, $\boldsymbol{x}^0 \equiv$ motion-uncompensated reconstruction, $\boldsymbol{\alpha}^0 \equiv$ observed dynamic region encoder, number of iterations n_{iter}.

Output: Reconstruction with region-based motion compensation, locally deformed regions, motion parameters.

for $i = 0 : n_{iter} - 1$

- $\boldsymbol{x}^{i+1} = \boldsymbol{x}^i - \gamma_{\boldsymbol{x}}^i \nabla_{\boldsymbol{x}} f\left(\boldsymbol{x}^i, \boldsymbol{\alpha}^i, \boldsymbol{p}^i\right)$
- $\boldsymbol{p}^{i+1} = \boldsymbol{p}^i - \gamma_{\boldsymbol{p}}^i \nabla_{\boldsymbol{p}} f\left(\boldsymbol{x}^i, \boldsymbol{\alpha}^i, \boldsymbol{p}^i\right)$
- $\boldsymbol{\alpha}^{i+1} = \boldsymbol{\alpha}^i - \gamma_{\boldsymbol{\alpha}}^i \nabla_{\boldsymbol{\alpha}} f\left(\boldsymbol{x}^i, \boldsymbol{\alpha}^i, \boldsymbol{p}^i\right)$
- Update the center of motion from $\text{Proj}_{\mathcal{S}}\left(\boldsymbol{\alpha}^{i+1}\right)$

The gradient of the objective function is analytically given by $\nabla f = \left[[\nabla_{\boldsymbol{x}} f]^T, [\nabla_{\boldsymbol{\alpha}} f]^T, [\nabla_{\boldsymbol{p}} f]^T\right]^T$, with

$$\nabla_{\boldsymbol{x}} f = \left\{ [(\boldsymbol{M}(\boldsymbol{p}) - \boldsymbol{I})\,\text{diag}\{\boldsymbol{\alpha}\} + \boldsymbol{I}] \underbrace{\begin{bmatrix} I \\ I \\ \vdots \\ I \end{bmatrix}}_{n \text{ blocks } I} \right\}^T \boldsymbol{W}^T \boldsymbol{r}, \tag{12}$$

$$\nabla_{\boldsymbol{\alpha}} f = \left\{ [\boldsymbol{M}(\boldsymbol{p}) - \boldsymbol{I}] \underbrace{\begin{bmatrix} \text{diag}\{\boldsymbol{x}\} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \text{diag}\{\boldsymbol{x}\} \end{bmatrix}}_{n \text{ blocks diag}\{\boldsymbol{x}\}} \right\}^T \boldsymbol{W}^T \boldsymbol{r}, \tag{13}$$

$$\nabla_{\boldsymbol{p}} f = [\nabla \boldsymbol{M}(\boldsymbol{p})(\boldsymbol{\alpha}[\circ]\boldsymbol{x})]^T \boldsymbol{W}^T \boldsymbol{r}, \tag{14}$$

where $\boldsymbol{r}$ is the residual of the system (5), computed as the follows:

$$\boldsymbol{r} = \boldsymbol{W}\left\{\overline{\boldsymbol{\alpha}}[\circ]\boldsymbol{x} + \boldsymbol{M}\left(\boldsymbol{p}\right)\left(\boldsymbol{\alpha}[\circ]\boldsymbol{x}\right)\right\} - \boldsymbol{b}. \tag{15}$$

The operators $\boldsymbol{M}\left(\boldsymbol{p}\right)$, $\boldsymbol{M}\left(\boldsymbol{p}\right)^T$ and $\nabla\boldsymbol{M}\left(\boldsymbol{p}\right)$ are all provided by a matrix-free and GPU-accelerated implementation of cubic image warping, its adjoint and its derivatives [11] designed to study continuous and differentiable affine motions. The operators $\boldsymbol{W}$ and $\boldsymbol{W}^T$ of the CT system are provided by the ASTRA Toolbox [12].

The objective function of the proposed method is non-convex towards the motion parameters $\boldsymbol{p}$. Nonetheless, the simultaneous convergence of the reconstruction and the parameter estimation schemes was empirically validated in [5]. The following property does not prove but supports the convergence of the region encoder estimation scheme when it is additionally combined with the reconstruction and the motion parameter estimation schemes, which shows the biconvexity of the objective function towards the reconstruction $\boldsymbol{x}$ and the region encoder $\boldsymbol{\alpha}$.

Theorem 1. *Let us assume the region encoder domain of the objective function f is extended to $[0,1]^{nN}$, f is then biconvex towards the reconstruction $\boldsymbol{x}$ and the region encoder $\boldsymbol{\alpha}$.*

Proof. The objective function (10) can be written as a quadratic form towards either the reconstruction variable $\boldsymbol{x}$ in the convex domain $[0,1]^n$ when $\boldsymbol{\alpha}$ and $\boldsymbol{p}$ are fixed:

$$f\big|_{\boldsymbol{\alpha},\boldsymbol{p}}(\boldsymbol{x}) = \frac{1}{2}\left\|\boldsymbol{W}\boldsymbol{P}\left(\boldsymbol{\alpha},\boldsymbol{p}\right)\boldsymbol{x} - \boldsymbol{b}\right\|_2^2, \tag{16}$$

with

$$\boldsymbol{P}\left(\boldsymbol{\alpha},\boldsymbol{p}\right) = \left[\left(\boldsymbol{M}(\boldsymbol{p}) - \boldsymbol{I}\right)\operatorname{diag}\left\{\boldsymbol{\alpha}\right\} + \boldsymbol{I}\right] \underbrace{\begin{bmatrix} I \\ I \\ \vdots \\ I \end{bmatrix}}_{n \text{ blocks } I}. \tag{17}$$

Similarly, in the extended convex domain $[0,1]^{nN}$ of $\boldsymbol{\alpha}$ when $\boldsymbol{x}$ and $\boldsymbol{p}$ are fixed, it yields:

$$f\big|_{\boldsymbol{x},\boldsymbol{p}}(\boldsymbol{\alpha}) = \frac{1}{2}\left\|\boldsymbol{W}\boldsymbol{Q}\left(\boldsymbol{x},\boldsymbol{p}\right)\boldsymbol{\alpha} - \boldsymbol{b}\right\|_2^2, \tag{18}$$

with

$$\boldsymbol{Q}\left(\boldsymbol{x},\boldsymbol{p}\right) = \left[\boldsymbol{M}(\boldsymbol{p}) - \boldsymbol{I}\right] \underbrace{\begin{bmatrix} \operatorname{diag}\left\{\boldsymbol{x}\right\} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \operatorname{diag}\left\{\boldsymbol{x}\right\} \end{bmatrix}}_{n \text{ blocks } \operatorname{diag}\left\{\boldsymbol{x}\right\}}. \tag{19}$$

Consequently, it is biconvex towards $\boldsymbol{x}$ and $\boldsymbol{\alpha}$ by [13]. □

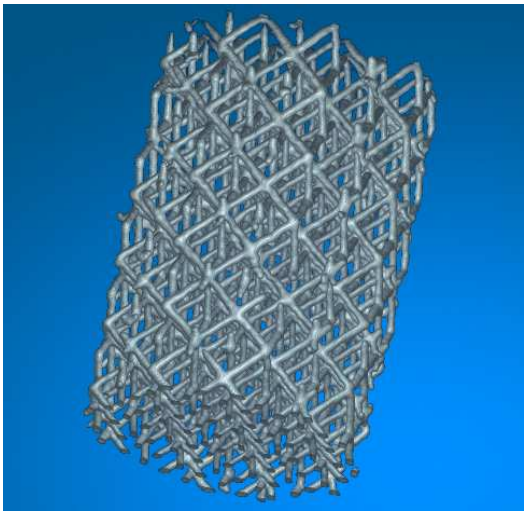
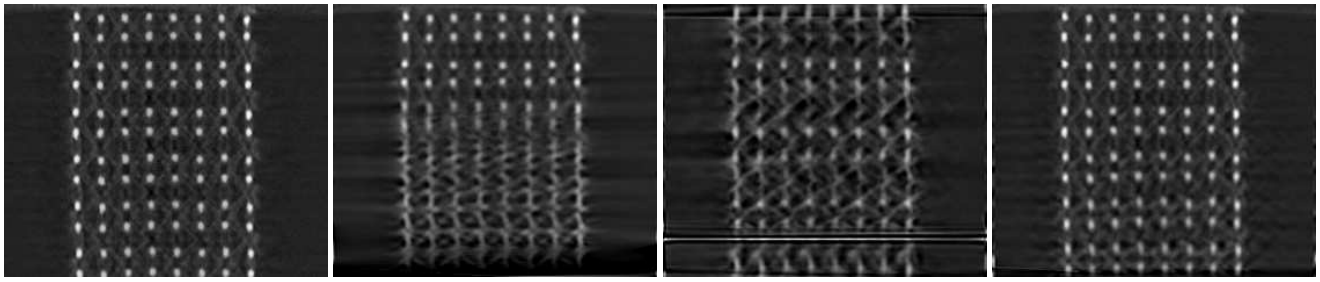


Fig. 1: The cylindrical bone scaffold.

3. EXPERIMENT AND RESULTS

We used a cylindrical bone scaffold of volume size $235 \times 280 \times 280$ (voxel) reconstructed from a real scan as the reference object (Fig. 1). Projection data was simulated by generating 720 uniformly-sampled cone beam projections spread over a full-rotation angular range. Gaussian noise with standard deviation of 1% of the peak gray value of the projection data was added to the sinogram. We assumed the object region from the top to the 50^{th} horizontal cross-section slice to be static in all angular projections. The projections were captured at discrete angular time points and the motion was simulated as continuous constant scaling in all three dimensions y-z, x-z and x-y respectively with the scaling factors range from 1 to 0.99, 0.99 and 1.25 in the deformed area. The experiment was considered in 5 subscans. The initial guess of the deformed region was the upper part of the object whose bottom z-coordinate is 55. At the i^{th} iteration, the stepsizes $\gamma_{\boldsymbol{x}}^i$ and $\gamma_{\boldsymbol{p}}^i$ were chosen following the Barzilai-Borwein formula [14] and the stepsize $\gamma_{\boldsymbol{\alpha}}^i$ was chosen constantly proportional to the quantity $1/i$. The z-coordinate of the center of motion was updated to be the z-coordinate of the bottom non-zero voxel of the intermediate estimated region encoder, when the x- and y- coordinates were in the center of the volume geometry. The chosen threshold value $\tau = 10^{-5}$ and the estimation of the deformed region was reduced to the estimation of the boundary between the static and deformed regions.

Convergence of the boundary estimation is achieved after around 15 iterations with a computation time of approximately 10 seconds per iteration. The reconstruction results are presented in Fig. 2, which shows a clear improvement over the reconstruction without motion compensation and the reconstruction wherein the deformation is supposed to appear in the entire volume of the object [5].



(a) Ground truth (b) without motion compensation (c) without region-based motion compensation [5] (d) with region-based motion compensation

Fig. 2: x-z cross-section of the reconstructions of the bone scaffold.

4. CONCLUSION AND FUTURE WORK

We have presented a reconstruction algorithm that combines accurate reconstruction, locally affine-deformed region identification and motion parameter estimation. Our method resulted in higher image quality compared to reconstruction that accounts for motion in the entire volume [5].

5. ACKNOWLEDGEMENT

This study is partially supported by the Research Foundation-Flanders (FWO) (SBO grant no. S007219N and PhD grant no. 1SA2920N). The authors would like to thank Prof. Martine Wevers and Dr. Jeroen Soete for sharing the reference object used for validating the method.

6. COMPLIANCE WITH ETHICAL STANDARDS

This is a numerical simulation study for which no ethical approval was required.

7. REFERENCES

[1] G. Zang, et al., "Space-time tomography for continuously deforming objects," *ACM Transactions on Graphics*, vol. 37, no. 4, pp. 1–14, 2018.

[2] G. Zang, et al., "Warp-and-project tomography for rapidly deforming objects," *ACM Transactions on Graphics*, vol. 38, no. 4, pp. 1–13, 2019.

[3] K. Ruymbeek and W. Vanroose, "Algorithm for the reconstruction of dynamic objects in CT-scanning using optical flow," *Journal of Computational and Applied Mathematics*, vol. 367, pp. 112459, 2020.

[4] M. Zehni, et al., "Joint angular refinement and reconstruction for single-particle Cryo-EM," *IEEE Transactions on Image Processing*, vol. 29, pp. 6151–6163, 2020.

[5] A.-T. Nguyen, et al., "An accelerated motion-compensated iterative reconstruction technique for dynamic computed tomography," in *Proc. SPIE 12242, Developments in X-Ray Tomography XIV, 122421F*, San Diego, CA, United States, 2022.

[6] J. Herrmann, et al., "Quantifying regional lung deformation using four-dimensional computed tomography: A comparison of conventional and oscillatory ventilation," *Frontiers in Physiology*, vol. 11, pp. 1–20, 2020.

[7] G. Van Eyndhoven, et al., "Region-based iterative reconstruction of structurally changing objects in CT," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 909–919, 2014.

[8] P. Gopal, et al., "Mitigating object prior-bias from sparse-projection tomographic reconstructions," *IEEE Transactions on Computational Imaging*, vol. 8, pp. 358–370, 2022.

[9] V. Slyusar, "A family of face products of matrices and its properties," *Cybernetics and Systems Analysis*, vol. 35, pp. 379–384, 1999.

[10] G. Van Eyndhoven, et al., "Combined motion estimation and reconstruction in tomography," in *Proceeding of the European Conference on Computer Vision (ECCV)*, 2012, pp. 12–21.

[11] J. Renders, et al., "ImWIP: image warping for inverse problems," Zenodo, DOI:10.5281/zenodo.5910755, 2023.

[12] W. van Aarle, et al., "Fast and flexible X-ray tomography using the ASTRA toolbox," *Optics Express*, vol. 24, no. 22, pp. 25129–25147, 2016.

[13] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[14] J. Barzilai and J. Borwein, "Two-point step size gradient methods," *IMA Journal of Numerical Analysis*, vol. 8, no. 1, pp. 141–148, 1988.