

Personalized X-Ray Reconstruction of the Proximal Femur via Intensity-Based Non-rigid 2D-3D Registration

Guoyan Zheng

Institute for Surgical Technology and Biomechanics,
University of Bern, CH-3014, Bern, Switzerland
Guoyan.Zheng@ieee.org

Abstract. This paper presents a new approach for reconstructing a patient-specific shape model and internal relative intensity distribution of the proximal femur from a limited number (e.g., 2) of calibrated C-arm images or X-ray radiographs. Our approach uses independent shape and appearance models that are learned from a set of training data to encode the *a priori* information about the proximal femur. An intensity-based non-rigid 2D-3D registration algorithm is then proposed to deformably fit the learned models to the input images. The fitting is conducted iteratively by minimizing the dissimilarity between the input images and the associated digitally reconstructed radiographs of the learned models together with regularization terms encoding the strain energy of the forward deformation and the smoothness of the inverse deformation. Comprehensive experiments conducted on images of cadaveric femurs and on clinical datasets demonstrate the efficacy of the present approach.

1 Introduction

Constructing a personalized three-dimensional (3D) model from a limited number of calibrated C-arm images or X-ray radiographs and a statistical model has drawn more and more attention. The reported techniques can be split into two main categories: those based on statistical models of shape [1][2][3] and those based on statistical models of shape and intensity [4][5][6][7][8][9]. The methods belonging to the former category typically require an implicit or explicit image segmentation which is error-prone and hard to achieve automatically. The errors in segmentation may lead to errors in the final reconstruction. In contrast, shape-intensity statistical model based methods directly compare the input reference images with the floating simulation images called digitally reconstructed radiographs (DRR), which are obtained by ray casting of a volume data instantiated from the learned model. No segmentation is required.

Most of the shape-intensity statistical model based methods are closely related with the Active Appearance Models (AAM) pioneered by Cootes et al. [10]. According to Matthews and Baker [11], there are basically two types of linear shape and appearance models, those model shape and appearance independently, and those which parameterize shape and appearance with a single

set of linear parameters. They referred the first set as *independent shape and appearance models* and the second as *combined shape and appearance models*. One good property associated with the independent formulation is that given a hypothesized shape the associated optimal appearance/intensity parameters may be determined analytically.

The contribution of this paper is a new approach for reconstructing a patient-specific shape model and internal relative intensity distribution of the proximal femur from sparse calibrated X-ray images. Unlike existing approaches, where they used either combined shape and appearance models [7] or a shape statistical model together with a density model approximated by Bernstein polynomials [4][6][8][9], our approach uses independent shape and appearance models that are learned from a set of training data. An intensity-based non-rigid 2D-3D registration algorithm is then proposed to deformably fit the learned models to the input images. The fitting is conducted iteratively by minimizing the dissimilarity between the input images and the associated DRRs of the learned models together with regularization terms encoding the strain energy of the forward deformation and the smoothness of the inverse deformation.

2 Construction of the Statistical Models

We constructed two independent shape and appearance models of the proximal femur, one for the left side and the other for the right side. Each independent shape and appearance model was constructed directly from CT data of 20 dry cadaveric proximal femurs of the associated side based on a two-stage procedure as described below.

Our two-stage independent shape and appearance model construction procedure follows the idea introduced by Rueckert et al. [12]. In the first stage, we chose one of the proximal femur volumes as the reference volume \mathbf{I}_0^{1st} and the diffeomorphic Demons algorithm [13] was used to establish the dense correspondences between the reference volume and each one of the 19 floating volumes. The output from the diffeomorphic Demons algorithm include: (1) the dense deformation fields $\{\mathbf{d}_i^{1st}; i = 1, \dots, 19\}$; and (2) the 19 non-rigidly deformed floating volumes $\{\mathbf{I}_i^{1st}; i = 1, \dots, 19\}$. Each deformation field \mathbf{d}_i^{1st} was expressed as a concatenation of 3D vectors describing the deformation at each voxel of the reference volume and each non-rigidly deformed floating volume \mathbf{I}_i^{1st} as a concatenation of gray values of each voxel in the reference volume. From these data, we computed the average deformation field $\bar{\mathbf{d}}^{1st} = (n^{-1}) \cdot \sum_{i=1}^n \mathbf{d}_i^{1st}$ and the average intensity distribution $\bar{\mathbf{I}}^{1st} = ((n+1)^{-1}) \cdot \sum_{i=0}^n \mathbf{I}_i^{1st}$.

The purpose of the second stage is to remove the possible bias introduced by the reference volume selection. To achieve this goal, we applied the average deformation field $\bar{\mathbf{d}}^{1st}$ to the reference volume \mathbf{I}_0^{1st} to create a new volume \mathbf{s}_0 and assigned the average intensity distribution $\bar{\mathbf{I}}^{1st}$ to this newly created volume. This new volume \mathbf{s}_0 was then used as the new reference volume in the second stage and all input 20 proximal femur volumes were regarded as the floating volumes. The diffeomorphic Demons algorithm [13] was used again to establish

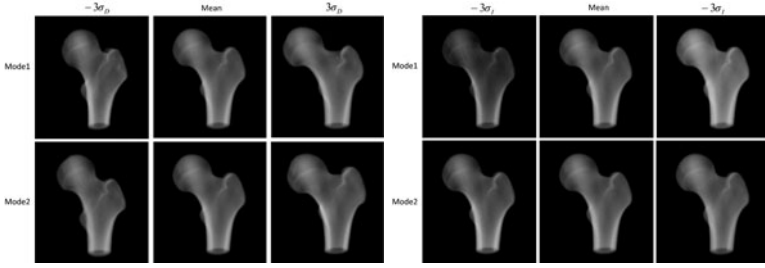


Fig. 1. Projections of the mean and the first two modes of variations of the shape (left) and the intensity (right) statistical models

the dense correspondences between the new reference volume and the other 20 floating volumes. We thus obtained a set of 20 new dense deformation field $\{\mathbf{d}_i; i = 1, \dots, 20\}$ and a set of 20 non-rigidly deformed floating volumes $\{\mathbf{I}_i; i = 1, \dots, 20\}$. We could then separately construct the shape statistical model and the intensity statistical model. The shape statistical model was constructed using following equations.

$$\begin{aligned}
 \mathbf{S}_D &= ((m - 1)^{-1}) \cdot \sum_{i=1}^m (\mathbf{d}_i - \bar{\mathbf{d}})(\mathbf{d}_i - \bar{\mathbf{d}})^T \\
 \bar{\mathbf{d}} &= (m^{-1}) \cdot \sum_{i=1}^m \mathbf{d}_i \\
 \mathbf{P}_D &= (\mathbf{p}_D^1, \mathbf{p}_D^2, \dots); \mathbf{S}_D \cdot \mathbf{p}_D^i = (\sigma_D^i)^2 \cdot \mathbf{p}_D^i \\
 \mathbf{d} &= \bar{\mathbf{d}} + \sum_{k=1}^{M_D} \alpha_D^k \sigma_D^k \mathbf{p}_D^k
 \end{aligned} \tag{1}$$

where $m = 20$ is the number of training samples; $\bar{\mathbf{d}}$ and \mathbf{S}_D are the average and the covariance matrix of the displacement fields, respectively; $\{(\sigma_D^i)^2\}$ and $\{\mathbf{p}_D^i\}$ are the eigen values and the eigen vectors of the shape statistical model, respectively; $\{\alpha_D^i\}$ are the model parameters; M_D is the cut-off points.

Fig. 1 shows the projections of the mean and the first two modes of variations of the shape (left) and the intensity (right) statistical models. Each instance of the shape statistical model was generated by evaluating $\mathbf{I}_s = \bar{\mathbf{I}}(\mathbf{s}_0 + \bar{\mathbf{d}} + \alpha_D \sigma_D^i \mathbf{p}_D^i)$, with $\alpha_D \in \{-3, 0, 3\}$. Each instance of the intensity statistical model was obtained by evaluating $\mathbf{I}_I = \bar{\mathbf{I}} + \alpha_I \sigma_I^i \mathbf{p}_I^i$, with $\alpha_I \in \{-3, 0, 3\}$, where $\{(\sigma_I^i)^2\}$ and $\{\mathbf{p}_I^i\}$ are the eigen values and the eigen vectors of the intensity statistical model, respectively; $\bar{\mathbf{I}} = (m^{-1}) \cdot \sum_{i=1}^m \mathbf{I}_i$ is the average intensity distribution.

From the instances of the shape statistical model (Fig. 1, left), it can be observed that although the instantiated intensity model is the same (here it is the average intensity distribution $\bar{\mathbf{I}}$), the change of shapes already encode the change of the relative intensity distribution inside the instantiated shape model (e.g., the thickness of the cortical bone along the shaft is different across the instantiated volumes). Thus, we argue that it is sufficient to search only in the space of shapes in order to estimate the personalized shape model and its internal relative intensity distribution. Based on this argument, we developed an intensity-based non-rigid 2D-3D registration algorithm as presented below.

3 Intensity-based Non-rigid 2D-3D Registration

In this work, we assume that we have a set of $Q \geq 2$ X-ray images and that all images are calibrated and co-registered to a common coordinate system called \mathbf{c} . Given an initial estimation of the registration parameters, our algorithm iteratively generates a volume and update the parameter estimation by minimizing the dissimilarity between the input images and the associated DRRs that are created from the instantiated volume.

Volume Instantiation and Alignment. The volume instantiation and alignment process is parameterized by two sets of parameters, i.e., the set of shape parameters $\mathbf{b} = (\alpha_D^1, \alpha_D^2, \dots, \alpha_D^{M_D})^T$ determining the forward deformation from the reference volume \mathbf{s}_0 to an instantiated volume \mathbf{s} and the set of parameters $\mathbf{a} = (\Lambda_x, \Lambda_y, \Lambda_z, \beta, \gamma, \theta, t_x, t_y, t_z)^T$ determining the scaled rigid transformation from the space of the instantiated volume \mathbf{s} to the common coordinate system \mathbf{c} , where the first three are scaling parameters; the middle three are rotational parameters and the last three are translational parameters. An instantiated volume that is aligned to the common coordinate system \mathbf{c} is defined by following equation:

$$\bar{\mathbf{I}}(x_c(\mathbf{a}, \mathbf{b})) = \bar{\mathbf{I}}(\mathbf{A}(\mathbf{a}) \circ \mathbf{W}(\mathbf{b}) \circ x_0) \tag{2}$$

where $\mathbf{A}(\mathbf{a})$ is the scaled rigid transformation and $\mathbf{W}(\mathbf{b})$ is the forward deformation.

Eq. (2) describes a forward warping that should be interpreted as follows. Given a voxel x_0 in \mathbf{s}_0 , the destination of this voxel under the forward transformation is $x_c(\mathbf{a}, \mathbf{b}) = \mathbf{A}(\mathbf{a}) \circ \mathbf{W}(\mathbf{b}) \circ x_0$. The aligned instance at voxel $x_c(\mathbf{a}, \mathbf{b})$ is set to the intensity $\bar{\mathbf{I}}(x_0)$, which then allows creating DRRs by simulating X-ray projection.

Implementing this forward warping to generate the model instance without holes is tricky and is best performed by backward warping. More specifically, as all the input images are calibrated, we can compute a back-projection ray for each pixel in the input images. We then do a sampling to get a set of discrete points along each ray. The intensity values at these points are then obtained by the backward warping with following inverse transformation:

$$x_0 = (\mathbf{W}(\mathbf{b}))^{-1} \circ ((\mathbf{A}(\mathbf{a}))^{-1} \circ x_c(\mathbf{a}, \mathbf{b})) \tag{3}$$

It is straightforward to compute the inverse of the scaled rigid transformation $\mathbf{A}(\mathbf{a})$. However, it is tricky to compute the inverse of the forward deformation $\mathbf{W}(\mathbf{b})$. In this paper, given an estimation of the shape parameters $\mathbf{b} = (\alpha_D^1, \alpha_D^2, \dots, \alpha_D^{M_D})^T$, we use the fixed-point approach as proposed by Chen et al. [14] to invert the forward deformation field.

Energy formulation. Given an instantiated volume, the two sets of parameters are updated iteratively by minimizing following energy function:

$$E(\mathbf{a}, \mathbf{b}) = E_{Image}(\mathbf{a}, \mathbf{b}) + \rho_1 \cdot \sum_{k=1}^{M_D} (\alpha_D^k)^2 (\sigma_D^k)^2 + \rho_2 \cdot \int_{\mathbf{s}} \|\nabla((\mathbf{W}(\mathbf{b}))^{-1})\|_2 dx \tag{4}$$

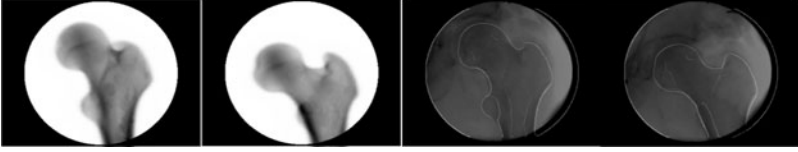


Fig. 2. Validation on a clinical dataset. The left two images show the projections of the reconstructed volume while the right two images show the original C-arm images superimposed by the contours detected from the left two images.

where the first term is the image dissimilarity energy term; the second term is the strain energy of the forward deformation, which, according to Cootes and Taylor [15], can be used to enforce the prior probabilities on the shape's deformations; the last term is a diffusion-like regularization term; ρ_1 and ρ_2 are regularization parameters.

We have chosen to use the robust dissimilarity measure that we introduced in [16] to compare the floating DRRs to the associated reference X-ray images. This dissimilarity measure is defined by modeling the q th observed difference image as a Markov random field with respect to the r th order neighborhood system $N = \{N_{i,j}^r\}$ and is described by following equation:

$$E_{Image}(\mathbf{a}, \mathbf{b}) = \sum_{q=1}^Q [\lambda \sum_{i,j}^{I,J} D_{q;(i,j)}^2(\mathbf{a}, \mathbf{b}) + (1 - \lambda) \sum_{i,j}^{I,J} \frac{1}{\text{card}(N_{i,j}^r)} \sum_{(i',j') \in N_{i,j}^r} (D_{q;(i,j)}(\mathbf{a}, \mathbf{b}) - D_{q;(i',j')}(\mathbf{a}, \mathbf{b}))^2] \quad (5)$$

where $I \times J$ is the size of each X-ray image; $D_q = \{D_{q;(i,j)}\}$ is the q th observed difference image. We refer interesting readers to [16] for the details about how the difference images are computed and about the details of above equation. The advantage of using such an energy function is that it has a least-squares form and can be effectively minimized by a Levenberg-Marquardt optimizer.

The registration Algorithm. Considering the fact that regularization on a vector field can be implemented as a kernel convolution [18], we developed following intensity-based non-rigid 2D-3D registration algorithm.

Algorithm (Intensity-based non-rigid 2D-3D Registration). The following two stages are executed until the convergence of the algorithm.

- Scaled rigid registration stage: The shape parameters are fixed to the current estimation \mathbf{b}_t and the Levenberg-Marquardt optimizer is used to iteratively minimize the image dissimilarity energy $E_{image}(\mathbf{a}, \mathbf{b}_t)$ in order to obtain a new estimation of the scaled rigid transformation parameters \mathbf{a}_{t+1} .
- Non-rigid registration stage: The scaled rigid transformation parameters are fixed to \mathbf{a}_{t+1} and the Levenberg-Marquardt optimizer is used again to iteratively estimate the new shape deformation parameters \mathbf{b}_{t+1} . At each iteration, following two steps are performed.

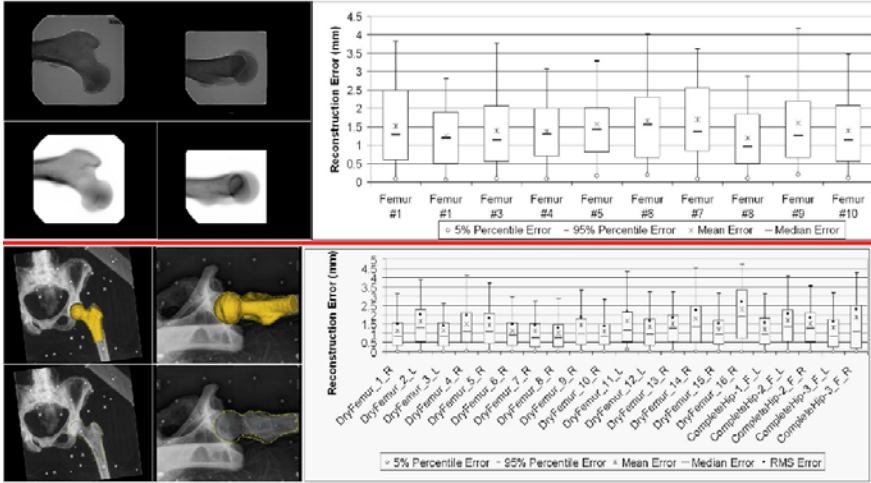


Fig. 3. Results of applying the present approach on sparse calibrated X-ray images. Top row: a reconstruction example as well as the errors of reconstructing all 10 femurs in Group I; bottom row: a reconstruction example as well as the errors of reconstructing all 21 femurs in Group II. In this example, the surface model (yellow) extracted from the reconstructed volume and its silhouettes are superimposed on the X-ray images.

- Step 1: Following the Levenberg-Marquardt optimizer, compute the gradient and the regularized Hessian of the first two terms of Eq. (4) with respect to the shape parameters, and then calculate an additive update $\Delta \mathbf{b}'_t$ of the shape parameters to get a new estimation $\mathbf{b}'_t = \mathbf{b}_t + \Delta \mathbf{b}'_t$.
- Step 2: Compute the forward deformation based on \mathbf{b}'_t and then invert the forward displacement fields using the fixed-point approach [14]. To implement the third term of Eq. (4), the inverse displacement fields are further regularized by applying a Gaussian kernel convolution. The regularized inverse displacement fields will then be used to instantiate a new volume to generate DRRs for the next iteration.

For all the experiments reported below, the scaled rigid transformation parameters \mathbf{a} were initialized with an anatomical landmarks based registration. All the shape parameters \mathbf{b} were initialized to zeros.

4 Experiments and Results

We conducted comprehensive experiments on 31 cadaveric femurs with different shape (none of them belongs to training population) and 3 clinical datasets to

validate the present approach. The 31 cadaveric femurs were divided into two groups. Group I consists of 10 left dry cadaveric femurs. For each femur in this group, we acquired two calibrated C-arm images around the proximal femur region. Group II contains the rest 21 femurs (7 left and 14 right) where 5 of them are connected with hips. For each femur in this group, we acquired two calibrated X-ray radiographs. Each one of the 3 clinical datasets contains two calibrated C-arm images of a patient. The cut-off point M_D was chosen to be 9 such that more than 90% of the total amount of variations is explained. The registration algorithm was implemented with CPU computation and no attempt was made to optimize its performance. Thus, one registration can take about 15 minutes. It can be accelerated by using modern GPU technology.

Experiment on 3 Clinical Datasets. In this experiment, due to the lack of ground truth, we mainly used these datasets to demonstrate qualitatively the performance of the present approach in clinical settings. Fig. 2 shows a reconstruction example when the present approach was applied to one of the three clinical datasets.

Experiment on 10 Cadaveric Femurs in Group I. In this experiment, the reconstruction accuracies were evaluated by randomly digitizing dozens points from the surface of each femur and then computing the distances from those digitized points to the associated surface model which was segmented from the reconstructed volume. A reconstruction example and the errors of reconstructing volumes of all 10 femurs in this group are shown in the top row of Fig.3. An average mean reconstruction accuracy of 1.5 mm was found.

Experiment on 21 Cadaveric Femurs in Group II. In this experiment, we used the method proposed by Aspert et al.[17] to compute the reconstruction errors, where the ground truth surface models were either obtained with a CT-scan reconstruction method or with a hand-held laser-scan reconstruction method (T-SCAN, Steinbichler, Neubeuern, Germany). The surface models segmented from the reconstructed volumes were then compared to the associated ground truth models to evaluate the reconstruction accuracies. A reconstruction example and the errors of reconstructing volumes of all 21 femurs in this group are shown in the bottom row of Fig. 3. An average mean reconstruction accuracy of 1.4 mm was found.

5 Conclusions

We presented a new approach to reconstruct a personalized shape model and internal relative intensity distribution of the proximal femur from sparse calibrated X-ray images. Our approach used independent shape and appearance models to encode the *a priori* information. An intensity-based non-rigid

2D-3D registration algorithm was then developed to fit the learned models to the input images. Results from our comprehensive experiments demonstrated its efficacy.

References

1. Fleute, M., Lavallée, S.: Nonrigid 3-D/2-D registration of images using statistical models. In: Taylor, C.J., Colchester, A.C.F. (eds.) MICCAI 1999. LNCS, vol. 1679, pp. 138–147. Springer, Heidelberg (1999)
2. Zheng, G., Gollmer, S., Schumann, S., Dong, X., Feilkas, F., Gonzalez Ballester, M.A.: A 2D/3D correspondence building method for reconstruction of a patient-specific 3D bone surface model using point distribution models and calibrated X-ray images. *Med. Image Anal.* 13, 883–899 (2009)
3. Baka, N., Niessen, W.J., Kaptein, B.L., van Walsum, T., Ferrarini, L., Reiber, J.H.C., Lelieveldt, B.P.F.: Correspondence free 3D statistical shape model fitting to sparse X-ray projections. In: Dawant, B.M., Haynor, D.R. (eds.) SPIE Medical Imaging 2010, vol. 7623, pp. 76230D-1–76230D9 (2010)
4. Sadowsky, O., Chintalapani, G., Taylor, R.H.: Deformable 2D-3D registration of the pelvis with a limited field of view, using shape statistics. In: Ayache, N., Ourselin, S., Maeder, A.J. (eds.) MICCAI 2007, Part II. LNCS, vol. 4792, pp. 519–526. Springer, Heidelberg (2007)
5. Hurvitz, A., Joskowicz, L.: Registration of a CT-like atlas to fluoroscopic X-ray images using intensity correspondences. *Int. J. CARS* 3, 493–504 (2008)
6. Chintalapani, G., Taylor, R.H.: Integrating statistical models of bone density into shape based 2D-3D registration framework. In: PMMIA 2009, pp. 1–11 (2009)
7. Humbert, L., Whitmarsh, T., De Craene, M., del Rio Barquero, L.M., Fritscher, K.D., Schubert, R., Eckstein, F., Link, T.M., Frangi, A.F.: 3D reconstruction of bone shape and bone mineral density distribution of the femur from DXA images. In: ISBI 2010, pp. 456–459. IEEE, Los Alamitos (2010)
8. Ahmad, O., Ramamurthi, K., Wilson, K.E., Engelke, K., Prince, R.L., Taylor, R.H.: Volumetric DXA (VXA) - A new method to extract 3D information from multiple in vivo DXA images. *J. Bone Miner. Res.* 25, 2468–2475 (2010)
9. Sadowsky, O., Lee, J., Sutter, E.G., Wall, S.J., Prince, J.L., Taylor, R.H.: Hybrid cone-beam tomographic reconstruction: incorporation of prior anatomical models to compensate for missing data. *IEEE T. Med. Imaging* 30, 69–83 (2011)
10. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. *IEEE T. Pattern Anal.* 23, 681–685 (2001)
11. Matthews, I., Baker, S.: Active appearance models revisited. *Int. J. Comput. Vision* 60, 135–164 (2004)
12. Rueckert, D., Frangi, A.F., Schnabel, J.A.: Automatic construction of 3D statistical deformation models using non-rigid registration. In: Niessen, W., Viergever, M. (eds.) MICCAI 2001. LNCS, vol. 2208, p. 77–84. Springer, Heidelberg (2001)
13. Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Non-parametric diffeomorphic image registration with the demons algorithm. In: Ayache, N., Ourselin, S., Maeder, A.J. (eds.) MICCAI 2007, Part II. LNCS, vol. 4792, pp. 319–326. Springer, Heidelberg (2007)

14. Chen, M., Lu, W., Chen, Q., Ruchala, K.J.: A simple fixed-point approach to invert a deformation field. *Med. Phys.* 35, 81–88 (2008)
15. Cootes, T.F., Taylor, C.J.: Combining point distribution models with shape models based on finite element analysis. *Image Vis. Computing* 13, 403–409 (1995)
16. Zheng, G.: Effective incorporating spatial information in a mutual information based 3D-2D registration of a CT volume to X-ray images. *Comput. Med. Imag. Grap.* 34, 553–562 (2010)
17. Aspert, N., Santa-Cruz, D., Ebrahimi, T.: Mesh: measuring error between surfaces using the hausdorff distance. In: *ICME 2002*, pp. 705–708. IEEE, Los Alamitos (2002)
18. Nielsen, M., Florack, L., Deriche, R.: Regularization, scale-space, and edge detection filters. *J. Math. Imaging Vis.* 7, 291–307 (1997)