Biomedical Imaging: A Computer Vision Perspective

Xiaoyi Jiang[1,2,3], Mohammad Dawood[1,2], Fabian Gigengack[1,2], Benjamin Risse[1,4], Sönke Schmid[1,2,3], Daniel Tenbrinck[1,2], and Klaus Schäfers[2,3]

[1] Department of Mathematics and Computer Science, University of Münster, Germany
[2] European Institute for Molecular Imaging (EIMI), University of Münster, Germany
[3] Cluster of Excellence EXC 1003, Cells in Motion, CiM, Münster, Germany
[4] Department of Neuro and Behavioral Biology, University of Münster, Germany

Abstract. Many computer vision algorithms have been successfully adapted and applied to biomedical imaging applications. However, biomedical computer vision is far beyond being only an application field. Indeed, it is a wide field with huge potential for developing novel concepts and algorithms and can be seen as a driving force for computer vision research. To emphasize this view of biomedical computer vision we consider a variety of important topics of biomedical imaging in this paper and exemplarily discuss some challenges, the related concepts, techniques, and algorithms.

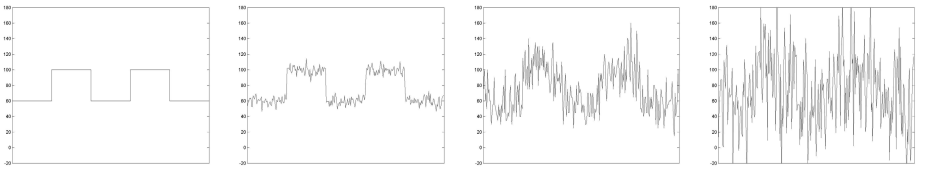
1 Introduction

The success story of modern biology and medicine is also one of imaging. It is the imaging techniques that enable biological experiments (for high-throughput behavioral screens or conformation analysis) and make the body of humans and animals anatomically or functionally visible for clinical purposes (medical procedures seeking to reveal, diagnose, or examine disease). With the widespread use of imaging modalities in fundamental research and routine clinical practice, researchers and physicians are faced with ever-increasing amount of image data to be analyzed and the quantitative outcomes of such analysis are getting increasingly important. Modern computer vision technology is thus indispensable to acquire and extract information out of the huge amount of data.

Computer vision has a long history and is becoming increasingly mature. Many computer vision algorithms have been successfully adapted and applied to biomedical imaging applications. However, biomedical imaging has several special characteristics which pose particular challenges, e.g.,

- Acquisition and enhancement techniques for challenging imaging situations are needed.
- The variety of different imaging sensors, each with its own physical principle and characteristics (e.g., noise modeling), often requires modality-specific treatment.

R. Wilson et al. (Eds.): CAIP 2013, Part I, LNCS 8047, pp. 1–19, 2013.



(a) (b) (c) (d)

Fig. 1. Illustration of three noise models. (a) Noise-free 1D signal. (b) Signal biased by additive Gaussian noise with $\sigma = 5$. (c) Signal biased by Poisson noise. (d) Signal biased by speckle noise with $\sigma = 5$ and $\gamma = 1$. (from [42])

- It is not seldom that different modalities are involved. Thus, algorithms must be designed to cope with multiple modalities.
- Due to the high complexity of many biomedical image analysis tasks, semi-automatic processing may be unavoidable in some cases. The design of intelligent and user-friendly interactive tools is a challenging task.
- Also the different body organs may require specific treatment.

As an example, the influence of noise modeling is considered. The following noise models are popular:

- Additive Gaussian noise: $f = \mu + \nu$, where μ is the unbiased image intensity and ν is a Gaussian-distributed random variable with expectation 0 and variance σ^2.
- Poisson noise ("photon counting noise"): This type of noise is signal-dependent and appears in a wide class of real-life applications, e.g., in positron emission tomography and fluorescence microscopy.
- Speckle noise: $f = \mu + \nu\mu^{\gamma/2}$ occurs in ultrasound imaging and is of multiplicative nature. Its dependency on the unbiased image intensity μ is controlled by the parameter γ. ν is the same as for additive Gaussian noise.

To illustrate the different characteristics of these noise forms a synthetic 1D signal and its corrupted versions are shown in Figure 1. We can observe that for similar parameters, the appearance of signal-dependent Poisson and speckle noise is in general stronger compared to the additive Gaussian noise. Their processing is thus definitively challenging and pushes the need for accurate data modeling in computer vision.

On the other hand, the special characteristics of biomedical imaging also give extra power to computer vision research. Multimodality can be helpful since they carry complementary information and their combined use may ease some image analysis tasks (e.g., segmentation [25]). Generally, a lot of knowledge specific to a particular application or object type may exist that should be accurately modeled and integrated into algorithms for dedicated processing towards improved performance.

Given the challenges discussed above, biomedical computer vision is far beyond simply adapting and applying advanced computer vision techniques to solve

real problems. It is also a wide field with huge potential of developing novel concepts, techniques, and algorithms. Indeed, biomedical imaging can be seen as a driving force for computer vision research.

In this paper this view of biomedical computer vision is emphasized by considering important topics of biomedical imaging: Minimum-cost boundary detection, region-based image segmentation, image registration, optical flow computation, and imaging techniques. Our intention is not to give a complete coverage of these topics, but rather exemplarily focus on typical challenges and the related concepts, techniques, and algorithms. The majority of the given examples is based on our own research and experiences in the respective fields.

2 Minimum-cost Boundary Detection

Quantification is one of the key words in biomedical imaging and requires robust, fast, and possibly automatic image segmentation algorithms. It can be either in the form of boundary detection or alternatively region-based segmentation. Automatic segmentation enables assessment of meaningful parameters, e.g., for diagnosis of pathological findings in clinical environments.

2.1 Live-Wire Techniques

Several paradigms of minimum-cost boundary detection exist in the literature. Among them the live-wire approach was initially introduced by Mortensen *et al.* [37] and Udupa *et al.* [51]. The user interactively picks a seed point on the boundary. Then, a live-wire is displayed in real time from the initial point to any subsequent position taken by the cursor. The entire 2D boundary is specified by means of a set of live-wire segments in this manner. The detection of segments is formulated as a graph searching problem, which finds the globally optimal (minimum-cost) path between an initial start pixel and an end pixel.

Placing seed points precisely on an object boundary may be difficult and tedious. To facilitate seed point placement, a cursor snap mechanism forces the mouse point to the pixel of maximum gradient magnitude within a user-specified neighborhood. The user-friendliness can be further increased by the live lane approach [20]. The user selects only the initial point. Subsequent points are selected automatically as the cursor is moved within a lane surrounding the boundary whose width changes as a function of the speed and acceleration of cursor motion.

Live-wire boundaries are piecewise optimal (between two seed points) and thus provide a balance between global optimality and local control. In contrast to statistical deformable approaches (e.g., [10,11,28]) no training is required. This semi-automatic technique has established itself as a robust and user-friendly method for the extraction of structure outlines for many biomedical applications.

A very fast implementation called live-wire on the fly is described in [19] which avoids unnecessary minimum-cost path computation during segmentation. Another important extension is the 3D generalization proposed in [18] to segment

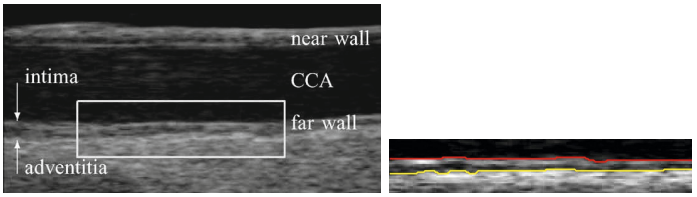


Fig. 2. B-mode CCA image (left) and detected intima and adventitia layer of far wall (right)

3D volume data or time sequences of 2D images. The key idea there is that the user specifies contours via live-wiring on a few slices that are *orthogonal* to the natural slices of the original data. If these slices are selected strategically, then one obtains a sufficient number of seed points in each natural slice which enable a subsequent automatic optimal boundary detection therein.

Live-wire techniques are a good example of designing intelligent and user-friendly interactive segmentation tools. They help to solve complex segmentation tasks by locally and non-extensively integrating the expertises and wishes of domain experts, which in turn also increases the user's faith in the automatic solution.

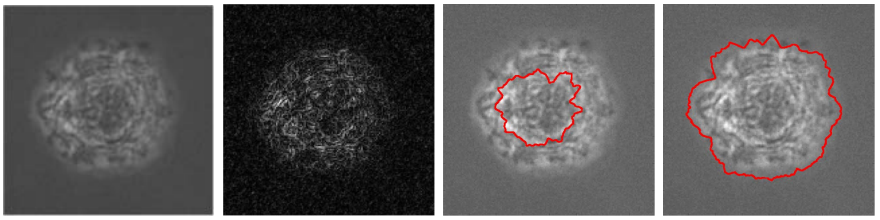
2.2 Dynamic Programming Based Boundary Detection

Dynamic programming (DP) is a popular technique for boundary detection due to its elegance, efficiency, and guarantee of optimality. One class of detectable boundaries starts from the left, passes each image column exactly once, and ends in the last column. An example is shown in Figure 2 for detecting the intimal and adventitial layers of the common carotid artery (CCA) in B-mode sonographic images [8]. Given an image of n rows and m columns, a total number of $O(n \cdot 3^{m-1})$ potential paths exist. However, the dynamic programming technique gives us an efficient algorithm for exactly finding the minimum-cost path with $O(mn)$ time and space [45].

Another, perhaps even more important, application class deals with closed boundaries. Based on a point p in the interior of the boundary, a polar transformation with p being the central point brings the original image into a matrix, in which a closed boundary becomes one from left to right afterwards. Finally, the detected boundary has to be transformed back to the original image space. This technique works well for star-shaped boundaries[1], particularly including (nearly) convex boundaries. Note that special care must be taken in order to guarantee the closedness of the detected boundary [47].

Typically, DP-based boundary detection assumes strong edges along the boundary and is thus based on gradient computation. In the simplest case the cost function is defined by the sum of gradient magnitudes. In practice, however,

[1] A star-shaped boundary is characterized by the existence of a point p such that for each interior point q the segment $\overline{pq}$ lies entirely inside.



(a) (b) (c) (d)

Fig. 3. (a) Tumor cell ROI; (b) gradient; (c) gradient-based optimal boundary; (d) region-based optimal contour (from [29])

gradient is not always a reliable measure to work with. One such example is the region-of-interest (ROI) of a tumor cell from microscopic imaging shown in Figure 3. Maximizing the sum of gradient magnitude does not produce satisfactory result.

There are only very few works on DP-based boundary detection using non-gradient information [35,53]. A challenge remains to develop boundary detection methods based on region information. A general framework for this purpose is proposed in [29]. A star-shaped contour C can be represented in polar form $r(\theta),\ \theta \in [0, 2\pi)$. Given the image boundary $B(\theta),\ \theta \in [0, 2\pi)$, the segmentation task can be generally formulated as one of optimizing the energy function:

$$E(C) \ = \ \int_0^{2\pi} \Big[\int_0^{r(\theta)} F_i(\theta, r) dr + \int_{r(\theta)}^{B(\theta)} F_o(\theta, r) dr \Big] d\theta \qquad (1)$$

Each region is assumed to be well represented by some model, which can be validated by a model testing function F_i (inside) and F_o (outside), respectively. This problem, however, cannot be solved by dynamic programming since the model parameters have to be estimated by the entirety of inside and outside of C. In [29] an approximation is thus made by *modeling each radial ray separately*, enabling to restrict the model testing functions $F_i(\theta, r)$ and $F_o(\theta, r)$ to a particular radial ray θ instead of the whole image. Then, a dynamic programming solution becomes possible for any representation model and model testing function independent of their form, complexity, and mathematical properties, e.g., differentiability. This universality gives the rather simple scheme of dynamic programming considerable power for real-world applications. In particular, robust estimation methods such as median-based approaches and L_1 norm (see Figure 3d for a related result) are highly desired for improved robustness. Also, sophisticated testing criteria like Fisher linear discriminant and others from machine learning theory provide extra useful options for measuring the separability of two distributions.

The principle of DP-based boundary detection can be extended in various ways. One extension is to simultaneously extract multiple boundaries [8,46], e.g., for detecting a pair of intimal and adventitial boundary in sonographic images (Figure 2). The domain of detectable boundaries can be further enlarged to

contain non-star-shaped objects. One such attempt from [30] allows the user to interactively specify and edit the general shape of the desired object by using a so-called rack, which basically corresponds to the object skeleton. The straightforward extension of the boundary class considered here to 3D is the terrain-like surface $z = f(x, y)$ (height field or discrete Monge surface). Unfortunately, there is no way of extending the dynamic programming solution to the 3D minimum-cost surface detection problem in an efficient manner. An optimal 3D graph search algorithm approach is presented in [32] with low polynomial time for this purpose. Similar to handling closed boundaries, cylindrical (tube-like) surfaces can be handled by first unfolding into a terrain-like surface using cylindrical coordinate transform. In addition to detecting minimum-cost surfaces this algorithm can also be applied to sequences of 2D images for temporally consistent boundary detection.

In practice, fast and easy-to-use algorithms like DP-based boundary detection are highly desired. To cite the biologist colleague who provided us the microscopic images used in [29] (see Figure 3): "I have literally tens of thousands of images per experiment" that must be processed within reasonable time. Therefore, further developments like boundary detection based on region information will have high practical impact.

3 Region-Based Image Segmentation

Region-based image segmentation is one of the fundamental problems in biomedical imaging for quantitative reasoning and diagnostic. Recently, mathematical tools such as level sets and variational methods led to significant improvements in image segmentation. However, a majority of works on image segmentation implicitly assume the given image to be biased by additive Gaussian noise, for instance the popular Mumford-Shah model [38]. Generally, it still lacks mature treatment of segmenting images with non-Gaussian noise models.

3.1 Discriminant Analysis Based Level Set Segmentation

The popular Chan-Vese (CV) approach [7], which is a special case of the Mumford-Shah formulation, uses a closed contour $\Gamma \subset \Omega$ to separate a given image domain Ω into two regions Ω_1, Ω_2. In particular, Γ is implicitly represented by the level sets of a Lipschitz function $\Phi\colon \Omega \to \Re$, i.e., $\Phi(x) < 0$ for $x \in \Omega_1$, $\Phi(x) = 0$ for $x \in \Gamma$, and $\Phi(x) > 0$ for $x \in \Omega_2$. Disregarding regularization of the segmentation area, the CV energy functional is given as:

$$\begin{aligned} E_{CV}(c_1, c_2, \Phi) \;&=\; \beta \int_{\Omega} \delta_0(\Phi(x))\, |\nabla \Phi(x)|\, dx \\ &+\; \lambda_1 \int_{\Omega} (c_1 \,-\, f(x))^2\, H(\Phi)\, dx \;+\; \lambda_2 \int_{\Omega} (c_2 \,-\, f(x))^2\, (1 - H(\Phi(x)))\, dx \end{aligned} \tag{2}$$

Here f is the perturbed image to be segmented and c_1 and c_2 are constant approximations of f in Ω_1 and Ω_2, respectively. The Heavyside function H is

used as indicator function for Ω_1, while δ_0 denotes the one-dimensional δ-Dirac measure.

In case of additive Gaussian noise (cf. Figure 1b) it is shown in [48] that for fixed c_1, c_2, the energy in Eq. (2) gets minimal if Φ partitions the data according to a natural *threshold* $t_{CV} = (c_1 + c_2)/2$ as in clustering where c_1 and c_2 are cluster centers. However, the situation in presence of multiplicative noise is different and the optimal threshold cannot be t_{CV} in this case (see [42] for details). In [48] a discriminant analysis (corresponding to the popular Otsu thresholding method) is thus applied to determine an optimal threshold t_O. Then, a new variational segmentation model is formulated as:

$$E(\Phi) \;=\; \frac{1}{2}\int_\Omega \operatorname{sgn}(\Phi(x))\,(f(x) \;-\; t_O)\,dx \;+\; \beta\int_\Omega \delta_0(\Phi(x))\,|\nabla\Phi(x)|\,dx \qquad (3)$$

This approach has been demonstrated to be superior to the Chan-Vese formulation on real patient data from echocardiography, which are known to be perturbed by multiplicative speckle noise.

3.2 Variational Segmentation Framework Incorporating Physical Noise Models

Despite its high popularity the Mumford-Shah formulation has not yet been investigated in a more general context of explicit physical noise modeling. Indeed, only few publications considered the effect of a specific noise model on the results of image segmentation [9,36]. A lot of segmentation problems need a suitable noise model, e.g., positron emission tomography or medical ultrasound imaging. Especially for data with poor statistics, i.e., with a low signal-to-noise ratio, it is important to consider the impact of the present noise model on the segmentation process.

In [42] a general segmentation framework for different physical noise models is presented, which also allows the incorporation of a-priori knowledge by using different regularization terms. For the special case of two-phase segmentation problems, the image domain Ω is partitioned into a background and a target subregion Ω_1 and Ω_2, respectively. An indicator function χ is introduced such that $\chi(x) = 1$ if $x \in \Omega_1$ and 0 otherwise (comparable to the Heavyside function $H(\Phi)$ in Eq. (2)). The data fidelity functions are defined by the negative log-likelihood functions derived from Bayesian modeling:

$$D_i(f, u_i) \;=\; -\log p_i(f \mid u_i) \quad \text{for } i \in \{1, 2\} \qquad (4)$$

where u_i is a smooth function for each subregion, which is chosen according to the assumed noise model for the given data f. Then, the energy functional for the two-phase segmentation problem is formulated as:

$$\begin{aligned} E(u_1, u_2, \chi) \;=\; \int_\Omega \chi(x)\,D_1(f, u_1) + (1 \;-\; \chi(x))\,D_2(f, u_2)\,\mathrm{d}x \\ + \alpha_1 R_1(u_1) + \alpha_2 R_2(u_2) + \beta\mathcal{H}^{n-1}(\Gamma) \end{aligned} \qquad (5)$$

Here $\mathcal{H}^{n-1}(\Gamma)$ is the $(n-1)$-dimensional Hausdorff measure. The regularization terms R_1 and R_2 are used to incorporate a-priori knowledge about the expected unbiased signals, e.g., H^1 seminorm, Fisher information, or TV regularization.

The choice of the probability densities $p_i(f \mid u_i)$ for $i = 1, 2$ crucially depends on the image formation process and hence on the noise model assumed for the data f and the subregion Ω_i. This is the place where physical noise modeling comes into play. In [42] the cases of Poisson and multiplicative speckle noise (cf. Figure 1c and 1d, respectively) have been intensively discussed.

In [50] the influence of three different noise models is investigated using this variational segmentation framework. In particular, shape priors are integrated as regularization term to the framework. It is demonstrated that correct physical noise modeling is of high importance for the computation of accurate segmentation results both in low-level as well as high-level segmentation.

The two approaches discussed above are representative for a variety of segmentation algorithms which fully utilize the knowledge about the specific characteristics of the image data at hand. A better modeling is the prerequisite for improved segmentation accuracy and robustness. This is especially important in biomedical imaging due to the variety of imaging modalities.

4 Image Registration

Image registration [21,34] aims at geometrically aligning two images of the same scene, which may be taken at different times, from different viewpoints, and by different sensors. It is among the most important tasks of biomedical imaging in practice. Given a template image $\mathcal{T} : \Omega \to \Re$ and a reference image $\mathcal{R} : \Omega \to \Re$, where $\Omega \subset \Re^d$ is the image domain and d the dimension, the registration yields a transformation $y : \Re^d \to \Re^d$ representing point-to-point correspondences between $\mathcal{T}$ and $\mathcal{R}$. To find y, the following functional has to be minimized:

$$\min_y \mathcal{D}(\mathcal{M}(\mathcal{T}, y), \mathcal{R}) + \alpha \mathcal{S}(y) \tag{6}$$

Here, $\mathcal{D}$ denotes the distance functional and the $\mathcal{M}$ transformation model, and $\mathcal{S}$ is the regularization functional. $\mathcal{D}$ measures the dissimilarity between the transformed template image and the fixed reference image. If both images are of the same modality, the sum-of-squared differences (SSD) can be used as a distance functional $\mathcal{D}$. In case of multimodal image registration information-theoretic measures, in particular, mutual information, are popular [39].

The SSD and related dissimilarity measures implicitly assume the intensity constancy between the template and reference image. Thus, we solely search for the optimal geometric transformation. In medical imaging, however, this assumption is not always satisfied. Such a problem instance appears in the context of motion correction in positron emission tomography (PET) [14].

PET requires relatively long image acquisition times in the range of minutes. In thoracic PET both respiratory and cardiac motion lead to spatially blurred images. To reduce motion artifacts in PET, so-called gating based techniques

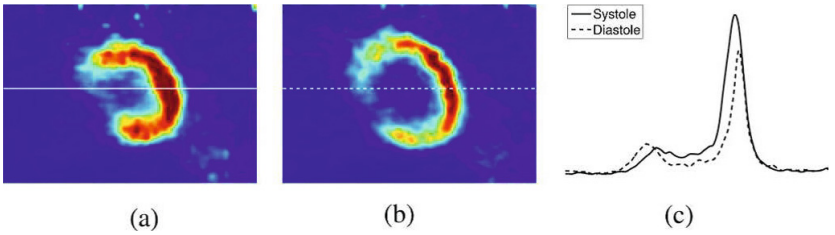


Fig. 4. Coronal slices of the left ventricle in a human heart during systole (a) and diastole (b) and corresponding line profiles (c) are shown for one patient. It can be observed that the maximum peaks in these line profiles vary a lot. (from [24])

were found useful, which decompose the whole dataset into parts that represent different breathing and/or cardiac phases [16]. After gating, each single gate shows less motion, however, suffers from a relatively low signal-to-noise ratio (SNR) as only a small portion of all available events is contained. After gating the data, each gate is reconstructed individually and registered to one assigned reference gate. The registered images are averaged afterwards to overcome the problem of low SNR. Tissue compression and the partial volume effect (PVE) lead to intensity modulations. Especially for relatively small structures like the myocardium the true uptake values are affected by the PVE. An example is given in Figure 4 where a systolic and diastolic slice (same respiratory phase) of a gated 3D dataset and line profiles are shown. Among others, the maximum intensity values of the two heart phases indicate that corresponding points can differ in intensity significantly.

In this situation an image registration mechanism is required which consists of simultaneous geometric transformation (spatially moving the pixels) and intensity modulation (redistributing the intensity values). In gating, all gates are formed over the same time interval. Hence, the total amount of radioactivity in each phase is approximately equal. In other words, in any respiratory and/or cardiac gate no radioactivity can be lost or added apart from some minor changes at the edges of the field of view. This property provides the foundation for a *mass-preserving image registration*. VAMPIRE (Variational Algorithm for Mass-Preserving Image REgistration) [24] incorporates a mass-preserving component by accounting for the volumetric change induced by the transformation y. Based on the integration by substitution theorem for multiple variables we have:

$$\int_{y(\Omega)} \mathcal{T}(x)dx \;=\; \int_{\Omega} \mathcal{T}(y(x))|\det(\nabla y(x))|dx \tag{7}$$

It guarantees the same total amount of radioactivity before and after applying the transformation y to $\mathcal{T}$. Therefore, for an image $\mathcal{T}$ and a transformation y, the mass-preserving transformation model is defined as:

$$\mathcal{M}^{\mathrm{MP}}(\mathcal{T}, y) \;:=\; \mathcal{T}(y) \cdot \det(\nabla y) \tag{8}$$

which is used in the registration functional (6) to enable simultaneous geometric transformation and intensity modulation.

In [24] this mass-preserving registration algorithm has been successfully applied to correct motion for dual – cardiac as well as respiratory – gated PET imaging. Motion estimation is also a fundamental requirement for super-resolution computation. More robust motion estimation based on mass-preserving registration thus facilitates improved super-resolution quality [52].

Similar to noise modeling discussed for region-based image segmentation, it is the explicit consideration of the mass-preserving property which enables improved image registration. This is another example of the benefit of accurate modeling in biomedical imaging.

5 Optical Flow Computation

Motion analysis is an important tool in biomedical imaging and optical flow estimation plays a central role in this context [2,22]. The basis of most optical flow algorithms is the brightness constancy[2]:

$$I(x, y, t) \;=\; I(x + u, y + v, t + 1) \tag{9}$$

which assumes that when a pixel moves from one image to another, its intensity (or color) does not change. In fact, this assumption combines a number of assumptions about the reflectance properties of the scene, the illumination in the scene, and the image formation process in the camera [2]. Linearizing this constancy equation by applying a first-order Taylor expansion to the right-hand side leads to the fundamental *optical flow constraint* (OFC):

$$u \cdot I_x + v \cdot I_y \;=\; -I_t \tag{10}$$

or more compactly:

$$f \cdot \nabla I \;=\; -I_t \tag{11}$$

with $f = (u, v)$ and $\nabla I = (I_x, I_y)$, which is used to derive optimization algorithms in a continuous setting.

In practice, however, this popular brightness constancy is not always valid. Other constancy terms have also been suggested including gradient, gradient magnitude, higher-order derivatives, e.g., on the (spatial) Hessian photometric invariants, texture features, and combination of multiple features (see [5] for a discussion). In the following two subsections we briefly discuss two additional variants from the medical imaging perspective.

5.1 Mass-Preserving Optical Flow

The problem of intensity modulations discussed for image registration in the previous section can also be tackled with optical flow techniques. According to

[2] For notation simplicity we consistently give the 2D version only. Its extension to n-D cases is straightforward.

the observation that the OFC is very similar to the continuity equation of fluid dynamics, Schunck [44] presented the *extended optical flow constraint* (EOFC):

$$f \cdot \nabla I + I \cdot \mathrm{div}(f) \;=\; -I_t \tag{12}$$

where $\mathrm{div}(f) = u_x + v_y$ is the divergence of f. Optical flow computaton based on EOFC has been studied and compared with ordinary OFC-based methods [4].

Interestingly, the EOFC has a physical interpretation of *mass preservation*. As shown by several researchers [3,12,40], this constraint is equivalent to a total brightness invariance hypothesis. The total brightness is defined as the sum of intensity values of a moving object. Instead of assuming that a point has a constant brightness over time, it is assumed that a moving object has a total brightness constant over time. Combined with a non-quadratic penalization a mass-preserving optical flow method has been applied for cardiac motion correction in 3D PET imaging [13]. In contrast to OFC-based optical flow [15,17], mass-preserving methods reflect better the physical reality of PET imaging.

Note that the idea behind the mass-preserving optical flow and the mass-preserving registration discussed in Section 4 is the same. Indeed, Eq. (12) can also be derived from Eq. (6), see [23]. Both registration and optical flow methods give us a powerful tool for solving mass-preserving motion estimation problems.

5.2 Histogram-Based Optical Flow

Multiplicative speckle noise (cf. Figure 1d) is characteristic for diagnostic ultrasound imaging. The origin of speckle are tiny inhomogeneities within the tissue, which reflect ultrasound waves but cannot be resolved by the ultrasound system. Speckle noise $f = \mu + \nu\mu^{\gamma/2}$ is of multiplicative nature, i.e., the noise variance directly depends on the underlying signal intensity.

The speckle noise has substantial impact on motion estimation. In fact, it turns out that the brightness constancy does not hold any more (see [49] for a mathematical proof). This can be demonstrated by the following simple experiment [49]. Starting from two pixel patches of size 5×5 with constant intensity values $\mu = 150$ and $\eta \in [0, 255]$, a realistic amount of speckle noise was added according to $f = \mu + \nu\mu^{\gamma/2}$ with $\gamma = 1.5$. The resulting pixel patches, denoted by X^{150} and Y^{η}, were compared pixelwise with the squared L_2-distance. Comparison of the two pixel patches was performed 10,000 times for every value of $\eta \in [0, 255]$. The simulation results (average distance of the two pixel patches and standard deviation) are plotted in Figure 5 (left). Normally, one would expect the minimum of the graph to be exactly at the value $\eta = \mu = 150$, i.e., both pixel patches have the same constant intensity before adding noise. However, the minimum of the graph is below the expected value. This discrepancy has been theoretically analyzed in [49], which predicts the minimum at $\eta \approx 141$ for the particular example as can be observed in Figure 5 (left).

In [49] it is argued that the overall distribution within a local image region remains approximately constant since the tissue characteristics remain and thus

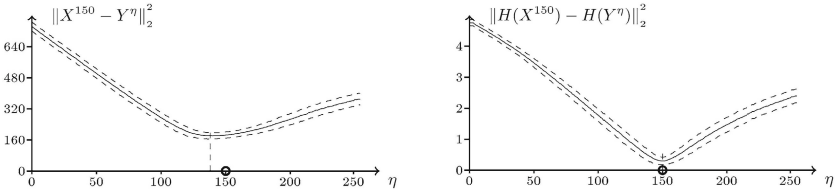


Fig. 5. Left: Average distance between two pixel patches biased by speckle noise. The global minimum is below the correct value of $\eta = 150$. Right: Average distance between the histograms of two pixel patches biased by speckle noise. The global minimum matches with the correct value of $\eta = 150$. In both case the two dashed lines represent the standard deviation of the 10,000 experiments. (from [49])

suggested to consider a small neighborhood around a pixel and compare the local statistics, i.e, local histograms as a discrete representation of the intensity distribution, of the images. This leads to the *histogram constancy constraint*:

$$H(x, y, t) \;=\; H(x + u, y + v, t + 1) \tag{13}$$

where H represents the cumulative histogram of the region surrounding the pixel (x, y) at time t. The validity of this new constraint has been mathematically proven in [49] and can also been seen in Figure 5 (right). On ultrasound data the derived histogram-based optical flow algorithm outperforms state-of-the-art general-purpose optical flow methods.

5.3 Periodic Optical Flow

In medical imaging some motion is inherently periodic. For example, this occurs in cardiac gated imaging, where images are obtained at different phases of the periodic cardiac cycle. Another example is in respiratory gated imaging, where the respiratory motion of the chest can also be described by a periodic model. Li and Yang [33] proposed optical flow estimation for a sequence of images wherein the inherent motion is periodic over time. Although in principle one could adopt a frame-by-frame approach to determine the motion fields, a joint estimation, in which all motion fields of a sequence are estimated simultaneously, explicitly exploits the inherent periodicity in image motion over time and can thus be advantageous against the framewise approach.

By applying Fourier series expansion, the components (u, v) at location (x, y) over time are modeled by:

$$u(x, y, t) = \sum_{l=1}^{L} \left[a_l(x, y) \cos \frac{2\pi l}{T} t + b_l(x, y) \sin \frac{2\pi l}{T} t \right] \tag{14}$$

$$v(x, y, t) = \sum_{l=1}^{L} \left[c_l(x, y) \cos \frac{2\pi l}{T} t + d_l(x, y) \sin \frac{2\pi l}{T} t \right] \tag{15}$$

where $a_l(x,y),\ b_l(x,y),\ c_l(x,y),\ d_l(x,y)$ are the coefficients associated with harmonic component l and L is the order of the harmonic representation. This motion model is embedded into the motion estimation for each pair of two successive images and the overall data term of the energy function to be minimized is the sum of all pairwise data terms from the brightness constancy.

While a number of constancy terms have been suggested in computer vision, the popular brightness constancy is dominating. The mass-preserving and histogram-based optical flow computation discussed above exemplarily demonstrate the need of finding suitable constancy terms in particular biomedical imaging scenarios. Periodic optical flow is a new concept and not fully explored yet. In both cases biomedical imaging provides large room for methodological development from a computer vision perspective.

6 Novel Imaging Techniques

Biomedical imaging has a broad range of subjects to be imaged and various imaging modalities. Despite of the enormous progresses there is still substantial room for further development of imaging technology. In the following we exemplarily describe two scenarios.

6.1 PET Imaging of Freely Moving Mice

In an on-going project we aim to track freely moving small animals with high precision inside a positron emission tomograph. Normally, the animals have to be anesthetized during 15-60 minutes of data acquisition to avoid motion artifacts. However, anesthesia influences the metabolism which is measured by PET. To avoid this, the aim of our project is to track awake and freely moving animals during the scan and use the information to correct the acquired PET data for motion. For this task a small animal chamber of $20\times10\times9$ cm was built (Figure 6) with a pair of stereo cameras positioned on both small sides of the chamber.

Due to the experimental setup highly distorted wide angle lenses have to be used. To reach the required tracking accuracy a high-precision lens distortion correction is crucial. First tests using a simple polynomial model for lens distortion correction lead to deviations from a pinhole camera model of up to 5 pixels. Therefore, more sophisticated methods are needed. Two high-precision lens distortion correction methods are described in [26,27]. In the latter case several images of a harp of wire are acquired and a massive amount of edge points is used to determine the parameters of a 11th grade polynomial distortion function. Both methods require a very accurately manufactured calibration pattern. In [43] another solution is suggested using a planar checkerboard pattern to provide very accurately detectable feature points even under distortion as a calibration pattern. Smoothed thin plate splines are applied to model the mapping between control points, leading to a mean accuracy below 0.084 pixel.

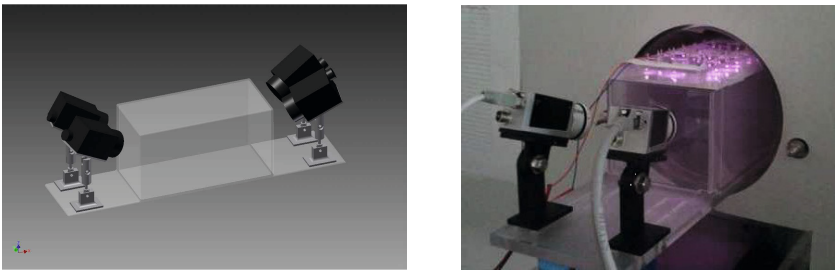


Fig. 6. Camera setup for PET imaging of freely moving mice. Left: Construction model of the animal chamber. Right: Manufactured chamber halfway inserted into a quad-HIDAC PET-scanner (16 cm in diameter). (from [43])

In addition to lens distortion correction we also need to solve other problems like feature detection and stereo vision in order to provide the technical fundament for a motion-corrected reconstruction. It is this bundle of computer vision solutions that will help to further improve the functionality of PET imaging towards imaging of freely moving mice.

6.2 FTIR-Based Imaging Method

The investigation of complex movement patterns of various organisms has become an integral subject of biological research. One of the most popular model organisms to study how the nervous system controls locomotion is Drosophila melanogaster (i.e., fruit fly). Drosophila is a holomethabolous insect. In the larval stage locomotion is confined to 2D, whereas the adult fly moves in two and three dimensions.

Work on freely flying fruit flies is still in its infancy because they form the so-called general multi-index assignment problem, which is nondeterministically polynomial-time hard (NP-hard) [6]. The current solutions are only able to track a small number of subjects for a short period [1,54]. From the computer vision perspective 3D tracking of flying flies is an unsolved challenge. Much efforts are still required to realize highly accurate tracking systems in order to fully meet the need of biological behavior studies.

In contrast, larval crawling occurs in two dimensions at relatively low speed. In principle, larval movement can be documented by a simple camera setup. However, recording of crawling larvae requires high contrast images, which are typically obtained by following sophisticated illumination protocols or dye applications [31]. For conventional, relatively low resolution tracking of larval locomotion, larvae are illuminated by incident or transmitted light and monitored by cameras with appropriate filters. This is technically challenging due to the semi-translucent body of these small animals. In addition, the observation of larvae is complicated by light reflections caused by the tracking surface. Thus, illumination problems aggravate faithful recordings of larval crawling paths and the poor signal to noise ratio complicates subsequent computer-based analysis.

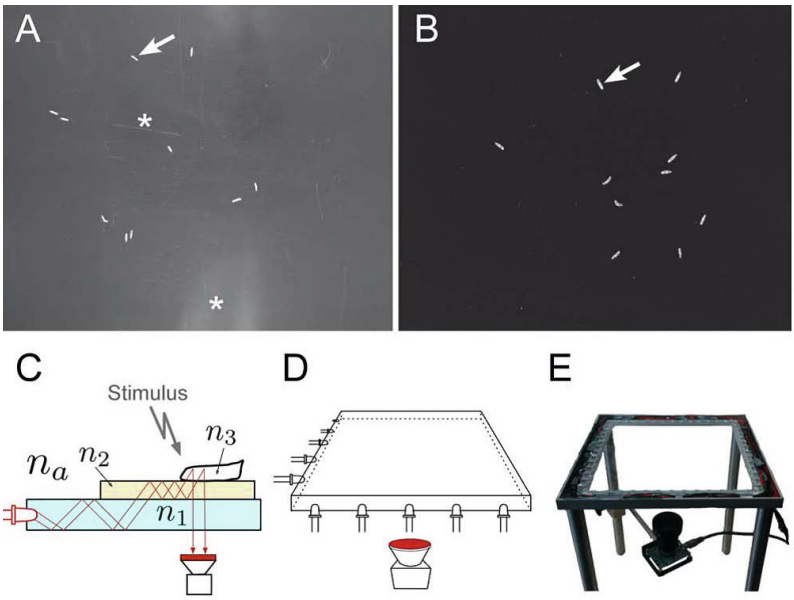


Fig. 7. The FIM setup. (A) Image of 10 larvae (arrow) imaged in a conventional setup. The asterisks denote scratches and reflections in the tracking surface. (B) Image of 10 larvae (arrow) imaged in the FIM setup with high contrast. (C) The principle of frustrated total internal reflection. n_a, n_1, n_2, and n_3 indicate different refractory indices of air, acrylic glass, agar and larvae respectively, an acrylic glass plate is flooded with infrared light (indicated by red lines). The camera is mounted below the tracking table. (D) Schematic drawing of the setup. (E) Image of the tracking table. (from [41])

A novel imaging technique based on frustrated total internal reflection (FTIR) is reported in [41], see Figure 7. Instead of directly illuminating crawling larvae, the frustrated total internal reflection is used to determine the contact surface between the animal and the substrate. In this FIM setup, an acrylic glass plate is flooded with infrared light. Due to the differences in the refractive indices of acrylic glass and air, it is completely reflected at the glass/air boundary (Figure 7C). To provide a moist crawling environment a thin agar layer is added. According to Snell's law, the light enters the agar layer since its refractive index (n_2) is higher than the refractive index of the acrylic glass (n_1). The larvae have an even higher optical density resulting in a higher refractive index (n_3), and thus, reflection is frustrated at the agar/larva interface and light enters the larval body. Here, light is reflected and since the reflection angle is smaller than the critical angle, the light passes through the different layers and can be detected by a camera equipped with an infrared filter (Figure 7D,E). This setup is easy to assemble and does not require cost-intensive equipment. This new imaging approach, named FIM (FTIR-based Imaging Method), provides an unprecedented high contrast view on crawling animals. Even without any background subtraction it generates constant image quality superior to previous setups. In addition,

it even allows to image internal organs. FIM is suitable for a wide range of biological applications and a wide range of organisms. Together with optimized tracking software it facilitates analysis of larval locomotion and will simplify genetic screening procedures.

7 Conclusion

Biomedical computer vision is far beyond simply adapting and applying advanced computer vision techniques to solve real problems. Biomedical imaging also poses new and challenging computer vision problems in order to cope with the complex and multifarious reality. In this paper we have exemplarily discussed a number of challenges and the related concepts and algorithms, mainly in the fields of our own research. They are well motivated by the practice. Biomedical imaging is full of such challenges and powerful computer vision solutions will immediately have benefit for the practice.

We need to understand how the domain experts work best with a technical system, which helps to design intelligent and user-friendly interactive tools. In addition, we are forced to have deeper understanding of the sources of signals and images to be processed, i.e., the objects of interest and biomedical devices. Only this way essential knowledge can be included for improved modeling and solution.

Many fundamental assumptions made when developing algorithms for biomedical imaging are shared by different - even non-biomedical - imaging modalities. For instance, the speckle noise model applies to both ultrasound and synthetic aperture radar imaging. Thus, the developed algorithms are of general interest and can be used in manifold application contexts.

Modern biology and medicine is a successful story of imaging. In the past biomedical computer vision has already established a vast body of powerful methods and tools. Continuous well-founded research will further enlarge the spectrum of successfully solved practical problems and thus continue to make a noticable contribution to biology and medicine.

Acknowledgments. The authors were supported by the Deutsche Forschungsgemeinschaft (DFG): SFB 656 MoBil (project B3, C3), EXC 1003 Cells in Motion – Cluster of Excellence, and DA 1064/3. Thanks go to Kristen Mills at Max Planck Institute for Intelligent Systems, Stuttgart, for providing the microscopic images.

References

1. Ardekani, R., Biyani, A., Dalton, J., Saltz, J., Arbeitman, M., Tower, J., Nuzhdin, S., Tavare, S.: Three-dimensional tracking and behaviour monitoring of multiple fruit flies. J. R. Soc. Interface 10(78), 20120547 (2013)
2. Baker, S., Scharstein, D., Lewis, J.P., Roth, S., Black, M.J., Szeliski, R.: A database and evaluation methodology for optical flow. International Journal of Computer Vision 92(1), 1–31 (2011)

3. Béréziat, D., Herlin, I., Younes, L.: A generalized optical flow constraint and its physical interpretation. In: Proc. of CVPR, pp. 487–492 (2000)
4. Bimbo, A.D., Nesi, P., Sanz, J.L.C.: Optical flow computation using extended constraints. IEEE Trans. on Image Processing 5(5), 720–739 (1996)
5. Bruhn, A.: Variational Optic Flow Computation – Accurate Modelling and Efficient Numerics. Ph.D. thesis, University of Saarland (2006)
6. Burkard, R., Dell'Amico, M., Martello, S.: Assignment Problems. Society for Industrial Mathematics (2009)
7. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Trans. on Image Processing 10(2), 266–277 (2001)
8. Cheng, D.C., Jiang, X.: Detections of arterial wall in sonographic artery images using dual dynamic programming. IEEE Trans. on Information Technology in Biomedicine 12(6), 792–799 (2008)
9. Chesnaud, C., Réfrégier, P., Boulet, V.: Statistical region snake-based segmentation adapted to different physical noise models. IEEE Trans. on Pattern Anaysis and Machine Intelligence 21(11), 1145–1157 (1999)
10. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. IEEE Trans. on Pattern Analysis and Machine Intelligence 23(6), 681–685 (2001)
11. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models-their training and application. Computer Vision and Image Understanding 61(1), 38–59 (1995)
12. Corpetti, T., Heitz, D., Arroyo, G., Memin, E., Santa-Cruz, A.: Fluid experimental flow estimation based on an optical-flow scheme. Experiments in Fluids 40(1), 80–97 (2006)
13. Dawood, M., Gigengack, F., Jiang, X., Schäfers, K.: A mass conservation-based optical flow method for cardiac motion correction in 3D-PET. Medical Physics 40(1), 012505 (2013)
14. Dawood, M., Jiang, X., Schäfers, K. (eds.): Correction Techniques in Emission Tomographic Imaging. CRC Press (2012)
15. Dawood, M., Büther, F., Jiang, X., Schäfers, K.P.: Respiratory motion correction in 3-D PET data with advanced optical flow algorithms. IEEE Trans. on Medical Imaging 27(8), 1164–1175 (2008)
16. Dawood, M., Büther, F., Stegger, L., Jiang, X., Schober, O., Schäfers, M., Schäfers, K.P.: Optimal number of respiratory gates in positron emission tomography: A cardiac patient study. Medical Physics 36(5), 1775–1784 (2009)
17. Dawood, M., Kösters, T., Fieseler, M., Büther, F., Jiang, X., Wübbeling, F., Schäfers, K.P.: Motion correction in respiratory gated cardiac PET/CT using multi-scale optical flow. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part II. LNCS, vol. 5242, pp. 155–162. Springer, Heidelberg (2008)
18. Falcão, A.X., Udupa, J.K.: A 3D generalization of user-steered live-wire segmentation. Medical Image Analysis 4(4), 389–402 (2000)
19. Falcão, A.X., Udupa, J.K., Miyazawa, F.K.: An ultra-fast user-steered image segementation paradigm: Live-wire-on-the-fly. IEEE Trans. on Medical Imaging 19(1), 55–62 (2000)
20. Falcão, A.X., Udupa, J.K., Samarasekera, S., Sharma, S., Hirsch, B.E., de Alencar Lotufo, R.: User-steered image segmentation paradigms: Live wire and live lane. Graphical Models and Image Processing 60(4), 233–260 (1998)
21. Fischer, B., Modersitzki, J.: Ill-posed medicine - an introduction to image registration. Inverse Problems 24(3), 034008 (2008)

22. Fleet, D., Weiss, Y.: Optical flow estimation. In: Paragios, N., Chen, Y., Fauregas, O. (eds.) The Handbook of Mathematical Models in Computer Vision, pp. 241–260. Springer (2005)
23. Gigengack, F.: Mass-Preserving Motion Correction and Multimodal Image Segementation in Positron Emission Tomography. Ph.D. thesis, University of Münster (2012)
24. Gigengack, F., Ruthotto, L., Burger, M., Wolters, C.H., Jiang, X., Schäfers, K.P.: Motion correction in dual gated cardiac PET using mass-preserving image registration. IEEE Trans. on Medical Imaging 31(3), 698–712 (2012)
25. Gigengack, F., Ruthotto, L., Jiang, X., Modersitzki, J., Burger, M., Hermann, S., Schäfers, K.P.: Atlas-based whole-body PET-CT segmentation using a passive contour distance. In: Menze, B.H., Langs, G., Lu, L., Montillo, A., Tu, Z., Criminisi, A. (eds.) MCV 2012. LNCS, vol. 7766, pp. 82–92. Springer, Heidelberg (2013)
26. von Gioi, R.G., Monasse, P., Morel, J.M., Tang, Z.: Towards high-precision lens distortion correction. In: Proc. of ICIP, pp. 4237–4240 (2010)
27. von Gioi, R.G., Monasse, P., Morel, J.M., Tang, Z.: Lens distortion correction with a calibration harp. In: Proc. of ICIP, pp. 617–620 (2011)
28. Heimann, T., Meinzer, H.P.: Statistical shape models for 3D medical image segmentation: A review. Medical Image Analysis 13(4), 543–563 (2009)
29. Jiang, X., Tenbrinck, D.: Region based contour detection by dynamic programming. In: Hancock, E., Smith, W., Wilson, R., Bors, A. (eds.) CAIP 2013, Part II. LNCS, vol. 8048, pp. 152–159. Springer, Heidelberg (2013)
30. Jiang, X., Große, A., Rothaus, K.: Interactive segmentation of non-star-shaped contours by dynamic programming. Pattern Recognition 44(9), 2008–2016 (2011)
31. Khurana, S., Atkinson, W.L.N.: Image enhancement for tracking the translucent larvae of drosophila melanogaster. PLoS ONE 5(12), e15259 (2010)
32. Li, K., Wu, X., Chen, D., Sonka, M.: Optimal surface segmentation in volumetric images - a graph-theoretic approach. IEEE Trans. on Pattern Analysis and Machine Intelligence 28(1), 119–134 (2006)
33. Li, L., Yang, Y.: Optical flow estimation for a periodic image sequence. IEEE Trans. on Image Processing 19(1), 1–10 (2010)
34. Maintz, J.B.A., Viergever, M.A.: A survey of medical image registration. Medical Image Analysis 2(1), 1–36 (1998)
35. Malon, C., Cosatto, E.: Dynamic radial contour extraction by splitting homogeneous areas. In: Real, P., Diaz-Pernil, D., Molina-Abril, H., Berciano, A., Kropatsch, W. (eds.) CAIP 2011, Part I. LNCS, vol. 6854, pp. 269–277. Springer, Heidelberg (2011)
36. Martin, P., Réfrégier, P., Goudail, F., Guérault, F.: Influence of the noise model on level set active contour segmentation. IEEE Trans. on Pattern Analysis and Machine Intelligence 26(6), 799–803 (2004)
37. Mortensen, E., Morse, B., Barrett, W.: Adaptive boundary detection using 'livewire' two-dimensional dynamic programming. In: IEEE Proc. Computers in Cardiology, pp. 635–638 (1992)
38. Mumford, D., Shah, J.: Optimal approximations by piecewise smooth functions and associated variational problems. Commun. Pure Appl. Math. 42, 577–685 (1989)
39. Pluim, J.P.W., Maintz, J.B.A., Viergever, M.A.: Mutual information based registration of medical images: A survey. IEEE Trans. on Medical Imaging 22(8), 986–1004 (2003)
40. Qiu, M.: Computing optical flow based on the mass-conserving assumption. In: Proc. of ICPR, pp. 7041–7044 (2000)

41. Risse, B., Thomas, S., Otto, N., Löpmeier, T., Valkov, D., Jiang, X., Klämbt, C.: FIM, a novel FTIR-based imaging method for high throughput locomotion analysis. PLoS ONE 8(1), e53963 (2013)
42. Sawatzky, A., Tenbrinck, D., Jiang, X., Burger, M.: A variational framework for region-based segmentation incorporating physical noise models. Journal of Mathematical Imaging and Vision (2013), doi:10.1007/s10851-013-0419-6
43. Schmid, S., Jiang, X., Schäfers, K.: High-precision lens distortion correction using smoothed thin plate splines. In: Hancock, E., Smith, W., Wilson, R., Bors, A. (eds.) CAIP 2013, Part II. LNCS, vol. 8048, pp. 432–439. Springer, Heidelberg (2013)
44. Schunck, B.: The motion constraint equation for optical flow. In: Proc. of ICPR, pp. 20–22 (1984)
45. Sonka, M., Hlavac, V., Boyle, R.: Image Processing, Analysis, and Machine Vision. Cengage Learning, 3rd edn. (2007)
46. Sun, C., Appleton, B.: Multiple paths extraction in images using a constrained expanded trellis. IEEE Trans. on Pattern Analysis and Machine Intelligence 27(12), 1923–1933 (2005)
47. Sun, C., Pallottino, S.: Circular shortest path in images. Pattern Recognition 36(3), 709–719 (2003)
48. Tenbrinck, D., Jiang, X.: Discriminant analysis based level set segmentation for ultrasound imaging. In: Hancock, E., Smith, W., Wilson, R., Bors, A. (eds.) CAIP 2013, Part II. LNCS, vol. 8048, pp. 144–151. Springer, Heidelberg (2013)
49. Tenbrinck, D., Schmid, S., Jiang, X., Schäfers, K., Stypmann, J.: Histogram-based optical flow for motion estimation in ultrasound imaging. Journal of Mathematical Imaging and Vision (2013), doi:10.1007/s10851-012-0398-z
50. Tenbrinck, D., Sawatzky, A., Jiang, X., Burger, M., Haffner, W., Willems, P., Paul, M., Stypmann, J.: Impact of physical noise modeling on image segmentation in echocardiography. In: Proc. of Eurographics Workshop on Visual Computing for Biomedicine, pp. 33–40 (2012)
51. Udupa, J., Samarasekera, S., Barrett, W.: Boundary detection via dynamic programming. In: Visualization in Biomedical Computing 1992, pp. 33–39 (1992)
52. Yan, H., Gigengack, F., Jiang, X., Schäfers, K.: Super-resolution in cardiac PET using mass-preserving image registration. In: Proc. of ICIP (2013)
53. Yu, M., Huang, Q., Jin, R., Song, E., Liu, H., Hung, C.C.: A novel segmentation method for convex lesions based on dynamic programming with local intra-class variance. In: Proc. of ACM Symposium on Applied Computing, pp. 39–44 (2012)
54. Zou, D., Zhao, Q., Wu, H.S., Chen, Y.Q.: Reconstructing 3d motion trajectories of particle swarms by global correspondence selection. In: Proc. of ICCV, pp. 1578–1585 (2009)