Questions and Answers

Chapter 8

Questions

(Q1) *What are photometric and geometric registration methods? How can these methods be compared?*

(Q2) *What is optical-flow?*

(Q3) *What is the aperture problem? How can it be solved?*

(Q4) *What are the advantages and drawbacks of optical-flow based registration?*

(Q5) *What are robust estimators?*

(Q6) *What are the advantages and drawbacks of robust M-estimators?*

(Q7) *How useful is a multiresolution scheme?*

(Q8) *How should the Gaussian standard deviation be chosen for building the multiresolution pyramid?*

(Q9) *What is a multigrid optimization scheme?*

(Q10) *What are the different options to regularize the deformation field?*

Answers

(A1) Registration methods can be broadly classified into two groups: geometric methods on the one hand that rely on the extraction and matching of sparse geometrical features; photometric methods on the other hand that use directly the luminance information. Geometrical methods make it possible to reduce the dimensionality of the problem. Furthermore, the registration is based on geometrical features that hopefully have an anatomical meaning.

Despite these advantages, the number of features that can be reproductively extracted are quite limited in the specific context of inter-subject brain matching. In addition to this, geometrical methods do not take into account all the information that is available. Therefore, photometric methods have been proposed. In the context of rigid registration, it has been shown that photometric methods are superior to geometric methods.

(A2) Optical flow is the assumption that the luminance of a given physical point does not change over the image sequence (sequence of images over time or over a database for instance). The brightness constancy assumption can be expressed as $f(s + \boldsymbol{w}_s, t_1) - f(s, t_2) = 0$, where s is a voxel of the volume, t_1 and t_2 are the indexes of the volumes to be registered, f is the luminance function and $\boldsymbol{w}$ the expected $3D$ displacement field. A first-order linear expansion of this equation may be preferred in case of small displacements: $\boldsymbol{\nabla} f(s, t) \cdot \boldsymbol{w}_s + f_t(s, t) = 0$ where $\boldsymbol{\nabla} f(s, t)$ stands for the spatial gradient of luminance and $f_t(s, t)$ is the voxelwise difference between the two volumes.

(A3) Because of the first-order linearization, it is possible to estimate only the projection of the deformation field on the local gradient. This is the classical aperture problem. To solve this problem, a regularization is explicitly introduced so as to smooth the deformation field and to constrain the estimation.

(A4) Thanks to the linearization, the optical flow can be estimated using fast techniques. Using locally parametric expressions for the deformation field, the optical flow leads to inverting linear systems. Contrary to multi-modal similarity measures such as mutual information, optical flow leads to fewer parameters and a faster implementation.

On the other hand, the optical flow is sensitive to luminance changes. In the context of inter-subject registration, the luminance conservation cannot stand. Two options can then be considered: either correct for intensity differences using histogram modeling and matching; or modify the estimation using robust functions to get rid of outliers.

(A5) Many computer vision problems amount to estimating hidden variables given observations. the modeling between variables and observations often treats observation data identically. It may be useful to neglect or reject some data because they correspond to noisy observations or because they correspond to an irrelevant secondary model. Therefore, robust estimators are used to get rid of inconsistent data.

Robust estimators can be characterized using three criteria: the relative efficiency that compares the variance of estimated parameters to the minimal variance (Cramer-Rao bound); the rupture point that is the maximum percentage of inconsistent data that the robust estimator can reject (the limit being 50%); and the complexity that traduces the additional computational complexity induced by robust estimators.

(A6) Robust M-estimators consist in replacing the quadratic cost by a cost function with known mathematical properties (see [Rousseewu 84] for complete details). It can be shown that robust M-estimators have a semi-quadratic formulation where an additional parameter is introduced. This additional parameter—comparable to a weight—tunes the rejection threshold. The minimization is then alternated: estimation of this additional parameter and minimization of the cost functional. Once the weight is estimated, the formulation is quadratic. This is the major benefit of robust M-estimators. The drawback is that robust M-estimators introduce an additional parameter that corresponds to the rejection threshold. The tuning o this parameter can be quite tedious and is application-specific.

(A7) A multiresolution scheme is quite classical in computer vision. A multiresolution pyramid is constructed by successive filtering and subsampling of the data. At each pyramid level, the dimensionality—and complexity—of the problem is reduced by a factor 2^3. In the context of image registration, the multiresolution scheme is useful to make the assumption of small deformations valid. As a matter of fact, at resolution level k, the magnitude of deformations is reduced by a factor 2^k.

(A8) To construct the multiresolution pyramid, successive filtering and subsampling of the data needs to be performed. Filtering are done using a Gaussian filter whose standard deviation is σ. The choice of σ is important.

Let us note r the voxel resolution in millimeter. If the volume is isotropic, the spectrum of the original signal is included in $]-F_E, F_E[$ with $F_E = 1/r$. Let us find a bound f_c for the filtered signal spectrum:

The frequency representation of the Gaussian filter is also a Gaussian distribution of standard deviation $\sigma' = 1/2\pi\sigma$. Since the convolution in the spatial domain amounts to a multiplication in the frequency domain, the kernel of the filtered signal is bounded by $3a'$ in the frequency domain.

To obey Shannon theorem, the sampling frequency should be at least twice the maximum signal frequency. We should have $F_E \geq 2f_c$ which gives:

$$\frac{1}{2r} \geq \frac{3 \times r}{2\pi\sigma} \Leftrightarrow \sigma \geq \frac{6r}{\pi}$$

(A9) Multigrid minimization schemes have been employed in computer vision when a large number of variables have to be estimated. This amounts to perform the estimation through nested subspaces. Multigrid methods are particularly useful when the cost functional are non-convex with a large number of local minima. The estimation is successively refined through nested subspaces whose dimension progressively increases. Multigrid techniques are employed so as to speed up the estimation process, especially when iterative solvers have to be used.

(A10) The regularization of the deformation field is necessary to obtain "reasonable" solutions. As a matter of fact, smooth deformations are often encountered. Two main options can be used: an intrinsic regularization or an explicit regularization.

An intrinsic regularization consists in using a regularized modeling of the deformation. This is for instance the case of Free-form deformation models or thin-plate-splines deformation models. The solution is sought in a subspace of smooth solutions.

An explicit regularization consists in expressing the regularization as a smoothness term. Classically, the cost functional is then composed of two terms: a data term (similarity) and a regularization term (smoothness of the deformation field). An alternative consists in alternating between maximizing the similarity and smoothing the field, like in the Demons' method.