In Search of Robustness and Efficiency via $\ell_1$— and $\ell_2$— Regularized Optimization for Physiological Motion Compensation

Angelica I. Aviles, Pilar Sobrevilla, Alicia Casals

Abstract—Compensating physiological motion in the context of minimally invasive cardiac surgery has become an attractive issue since it outperforms traditional cardiac procedures offering remarkable benefits. Owing to space restrictions, computer vision techniques have proven to be the most practical and suitable solution. However, the lack of robustness and efficiency of existing methods makes physiological motion compensation an open and challenging problem. This work focuses on increasing robustness and efficiency via exploration of the classes of $\ell_1$— and $\ell_2$—regularized optimization, emphasizing the use of explicit regularization. Both approaches are based on natural features of the heart using intensity information. Results pointed out the $\ell_1$—regularized optimization class as the best since it offered the shortest computational cost, the smallest average error and it proved to work even under complex deformations.

Keywords—Motion Compensation, Optimization, Regularization, Beating Heart Surgery, Ill-posed problem.

I. INTRODUCTION

MINIMALLY invasive beating heart surgery has demonstrated to provide relevant benefits in comparison with traditional cardiac surgery such as [1], [2]: smaller incisions; shorter term rehabilitation; less bleeding; faster recovery; cosmetic improvement; and reduction of cognitive and neurological effects. However, since the heart is beating while intervention is performed, the surgeon has to face two sources of disturbances: respiration and heartbeat. These physiological motions hinder surgeon’s gestures and limit precision during surgery. Thus, a main research goal is to compensate them in order to give to the surgeon the sensation of working in a motionless area.

The quasiperiodic behaviour of the heart; this, together with the high precision requirements, the need of operating in real time, the surface characteristics, the limited workspace, and illumination changes make physiological motion compensation be a real challenge. To cope with these drawbacks in some works (e.g. [3], [4], [5], [6]) the use of mechanical stabilization devices, which are positioned over the heart surface to keep the region of interest in steady state, are proposed. However, their performances have been reported to be unsatisfying because: physiological motion still needs to be manually compensated by the surgeons, devices can only be applied on the top of the heart surface, can damage the heart tissue, and biocompatibility studies are not carried out.

Moreover, the use of various sensors to compensate physiological motion has been investigated by some authors, including accelerometers [7], laser scan system [8], whisker sensor [9], an optical fiber sensor [10] and a sensor based on microwave doppler radar [11]. Besides the sensor size problem; a serious issues are: their long-term stability; their shape; and the biocompatibility (e.g. hemocomparibility). Because of previous drawbacks, the most practical and suitable solution is the use of a vision sensor, which avoids above-mentioned drawbacks since it is integrated on the endoscope. Thus, the use of Computer Vision Techniques (CVT) are adopted in this work.

The feasibility of CVT for physiological motion compensation was first explored by Nakamura et al. [12]. In that work, the authors introduced the concept of heart synchronization for minimally invasive cardiac surgery. Also, they developed a heartbeat synchronization system, in which artificial markers over the heart were used for tracking the region of interest. Artificial markers were also employed in [13], but afterwards avoid them using a texture-based approach. Similarity, passive colored markers attached to the heart, to do feature extraction, were adopted in [14], in which a physic-based heart motion tracking method was proposed. That is, authors described the heart surface as a physical elastic body in form of partial differential equations. However, the use of artificial markers is impractical because of the presence of liquids (e.g. blood) and vapors during the surgery, and the difficulty of putting them on the heart.

To avoid these problems Ortmayer et al. [15] explored natural landmarks on heart surface for estimating heart motion. In their work a reduced affine tracking method was proposed for tracking the 2D displacement of salient features on the heart surface. Likewise, a solution based on the combination of both thin-plate splines and second-order minimization is put forward in [16]. Another solution has been presented in [17], in which the SURF method is used in order to recover distinct regions that are used in the Lucas-Kanade tracker.

In this work, two approaches, which make use of the $\ell_1$— and $\ell_2$—regularized optimization classes are presented and compared, with the main aim of generating a solution capable of working in realistic clinical environments. Moreover, the introduction of explicit regularization is highlighted in order to solve the ill-posed problem, offering stability to the system and helping the optimization process via decreasing the number of
local minima.

II. PROBLEM STATEMENT

The setup for heart motion compensation is part of the Robotic Surgical System given in Fig. 1. Besides the motion compensation module, the system includes: a processing module for image correction and enhancement; a prediction module, which ensures system operation during occlusions (e.g. when the camera is blocked by an instrument); and a control module, which synchronizes the surgical instruments with the heart motion.

The solution for compensating physiological motion is formulated as an energy minimization problem, in which the available information is given by a vision sensor. Thus, let \( \mathbf{I}_t : \Omega_0 \rightarrow \mathbb{R} \) and \( \mathbf{I}_a : \Omega_0 \rightarrow \mathbb{R} \) be the fixed and the acquired images, respectively, given a each time instant, in which the transformation/displacement, \( u \), minimizes the total energy \( E_t \).

More specifically, the defined energy functional is given by two main terms: the dissimilarity term, \( E_d \), which evaluates the discrepancy between the two given images and the regularization term, \( E_r \), that allows fulfilling the Hadamard’s postulate [18] and generates a well-defined displacement field represented by \( u, \alpha \in \mathbb{R}^d \), gives a balance between both terms, and \( x \) represents a vector in \( \mathbb{R}^d \).

\[
\arg\min_u E_t(u) = E_d(\mathbf{I}_t(x), \mathbf{I}_a(u(x) + x)) + \alpha E_r(u(x)) \tag{1}
\]

A deformation model has to be carefully selected to parametrize \( u \), which has a direct effect over the global behaviour of the system. This model has also to be characterized by two main factors: short computational time and valuable information [19]. A classification of deformation models was proposed in [20] with three main categories i) physical models, ii) knowledge models and iii) approximation and interpolation models.

Although the first and second categories have been used in medical applications, their disadvantages are the complex computation and the deficiency in deformation retrieval. Therefore, the third category is the most adequate option due to its ability to deal with complex deformations, and to the convenient time consumption. In this category well-known models can be found such as thin-plate splines, elastic body splines and piecewise affine. Their main drawbacks are inverse inconsistency, lack of mathematical optimality and continuity, and high computational demand. An interesting option is B-splines that offer various desirable benefits [21], [22]: easy manipulation, compact support, low computational cost, optimal mathematical properties, multiresolution, Hölder continuous of order \( n \), affine invariance and preservation of both convexity and diminishing properties.

Based on previous arguments, B-splines have been selected in this work to parametrize the displacement field \( u \). Let \( C_j \) be all control points in \( d \)-dimensions belonging to a lattice as \( \Phi \) that allows fulfilling the parametrization of \( u \) using b-splines of degree 3 (i.e. cubic) is written as

\[
u_d(x) = \sum_{j_1=0}^{n} \cdots \sum_{j_d=0}^{n} \mathbf{C}_{j_1,\ldots,j_d} \prod_{k=1}^{d} \mathbf{Y}(u_{k}) \cdot \{x \in \mathbb{R}^d\}
\]

Defining the position vector \( x = (x,y) \) and the size of the lattice \( \Phi \), then, \( p = \frac{1}{\sqrt{2}} - 1 \) and \( q = \frac{1}{\sqrt{2}} - 1 \). Moreover, let \( Y_{k,m} \) be basis functions expressed as

\[
\begin{align*}
Y_0(\mu) &= (1 - \mu)^3/6 \\
Y_1(\mu) &= (4 + 3\mu^3 - 6\mu^2)/6 \\
Y_2(\mu) &= (1 - 3\mu^3 - 3\mu^2 + 3\mu)/6 \\
Y_3(\mu) &= \mu^3/6
\end{align*}
\]
III. PROPOSED SOLUTIONS

In this section, the generalized optimization problem presented in (1) is reformulated using the classes of $\ell_1$- and $\ell_2$-regularized optimizations. The general process that is carried out in both cases can be seen in Fig. 2. These formulations use intensity information of the images as matching criteria and two well-known regularization methods: Tikhonov [23] and Total Variation (TV) [24].

The idea of use explicit regularization in this application is: to guarantee an existing and stable solution via restriction of the space of solutions; to penalize oscillating deformations (i.e., generate a well-defined displacement field); and to decrease the computational cost of the optimization process. It is worth noticing that regularization is applied to the gradient of the displacement field, $\nabla u \in \mathbb{R}^2$, that is discretized as follows: Let $I_{roi} \in \Omega$ be a region of interest formed by $N \times N$ pixels with $i, j$ positions, then, $\nabla u(x) = u_{i+1,j} - u_{i,j} \parallel \nabla u(x) = u_{i,j+1} - u_{i,j}$ when $N < i \parallel N < j$, otherwise ($N = i \parallel N = j$) $\nabla u = 0$.

A. $\ell_1$-Based Formulation

Motivated by the fact that proposals of this problem presented up to now are oriented to use the norm $\ell_2$, exploration of the class of $\ell_1$-regularized optimization in the context of heart motion is proposed. This class offers different advantages such as: it works well with outliers; is related to robustness; outperforms solutions (in certain applications) obtained with the class $\ell_2$; and is non-sensitivity to changes of intensity.

In order to obtain the dissimilarity between the pixels in $I_f$ and $I_a$, the Sum of Absolute Difference (SAD) method is applied. Thus, rewriting first term, $E_d$, from (1), it results in

$$E_d = \frac{1}{2} \sum_{x \in \Omega} \left| (I_a(u(x) + x) - I_f(x)) \right| \ dx$$

$$= \frac{1}{2} \sum_{x \in \Omega} \left| (I_a(u(x) + x) - I_f(x)) \right| \tag{4}$$

The regularization term, $E_r$, is reformulated using TV method, in which penalization is carried out via the absolute gradient of the displacement field. This method has been chosen due to its simplicity, efficiency, easier interpretation and its ability to preserve discontinuities in the displacement field. Moreover, its desirable mathematical properties as lower semi-continuity, convexity and homogeneity [25]. Thus, reformulating second term, $E_r$, from (1) applying TV, it is given by

$$E_r = \sum_{d=1}^{N} \left\{ \frac{1}{2} \sum_{x \in \Omega} \left| \nabla u_d(x) \right| \ dx \right\}$$

$$+ \alpha \sum_{x \in \Omega} \left| (\nabla u_1(x), \nabla u_2(x)) \right| - \lambda \sum_{x \in \Omega} \log c_x(u) \right\} \tag{5}$$

The main inconvenience of the previous formulation is that it is non-differentiable at zero. As a consequence, the optimization process is hindered, and standard methods are not directly applicable. A classic alternative is to use subgradient methods [26] or $\epsilon$-subgradient methods [27], but they converge slowly and they have the lack of efficiency in certain cases. Another option is to treat the unconstrained non-differentiable problem as a constrained problem. The simplest and effective alternative is to redefine the functional with a differentiable approximation, allowing the use of classic unconstraint optimization methods. In this case, in order to avoid moving to the restricted area (i.e., zero), the log-barrier function given in (6) is applied, where $\log(\cdot)$ is the natural logarithm and $\lambda$ is the value that determines the impact of the barrier over the function.

$$f_{log}(u) = -\lambda \sum_{x \in \Omega} \log c_x(u)$$

$$\{ u \in \mathbb{R}^N | c_x(u) > 0 \text{ for all } x \in \Omega \} \tag{6}$$

$$\nabla f_{log}(u) = - \frac{\lambda}{c_x(u)} \nabla c_x(u)$$

Therefore, using (4), (5) and (6), the energy functional becomes differentiable as can be seen in next Equation

$$\arg \min_{u} \left\{ \frac{1}{2} \sum_{x \in \Omega} \left| (I_a(u(x) + x) - I_f(x)) \right| + \alpha \sum_{x \in \Omega} \left| (\nabla u_1(x), \nabla u_2(x)) \right| - \lambda \sum_{x \in \Omega} \log c_x(u) \right\}$$

$$(\lambda, \alpha, Ed, Er, \log) \tag{7}$$

where $c_x(u) = u^2$. Since, the resulting minimization problem described in (7) is a smooth optimization problem, it can be solved using standard unconstrained methods. In this case, the
Newton's method is applied, which allows having an excellent global convergence.

### B. $\ell^2$-Based Formulation

The class of $\ell^2$-regularized optimization is the most popular way to deal with ill-posed problems since the optimization process is straightforward. That is, the functional is differentiable and the derivatives are continuous. In the case of the first term of (1), $E_d$, reformulation is given using Sum of Squared Differences (SSD) method. Apart from its differentiability, its simplicity and low computational cost makes this method a good candidate to measure the discrepancy between the two given images, $I_r$ and $I_a$. Thus, rewriting first term from (1) using SSD, it becomes in

$$E_d = \frac{1}{2} \int_\Omega (I_r(x) - I_a(u(x) + x))^2 dx$$

It is worth noting that this method is sensitive to changes of intensities, hence, a normalization is advisable. The regularization term, $E_r$, is generated using Tikhonov method. Therefore, $E_r$ is given by

$$E_r = \frac{1}{2} \sum_{i=1}^6 \int_\Omega |\nabla u_i(x)|^2 dx$$

Then using (8) and (9) the total energy becomes

$$\arg \min_u \left\{ \frac{1}{2} \sum_{x \in \Omega} (I_r(x) - I_a(u(x) + x))^2 + \frac{\alpha^2}{2} \sum_{x \in \Omega} |\nabla u_1(x)|^2 + |\nabla u_2(x)|^2 \right\}$$

In order to minimize (10), it is necessary to choose an optimization method. Traditional methods for solving $\ell^2$ include Gradient Descent (GD) [28], Newton's method (NM), Nonlinear Conjugate Gradient [29] (NCG), Evolutionary Optimization Algorithms (EOA) [30] etc. However, they present some drawbacks. They can get stuck at local minima; sometimes need an infinite number of iterations to converge; or have a slow rate of convergence. A common alternative used in this context to avoid these issues is the well-known Levenberg-Marquardt (LM) [31], [32]. Although LM outperforms above inconvenients, in some cases it has to deal with the slow convergence problem. A more attractive option, which offers an excellent quality of results with better computational time than LM, is the Powell’s dogleg (DL) [33]. The idea behind this method is to take the best of different methods, that is, the convergence speed of the gauss-newton, the excellent global convergence of the gradient descent, handling them via an explicit trust-region. DL is used to minimize (10) in order to improve the optimization process. This method is chosen because of their characteristics, and because its performance has not been yet explored in this context.

### IV. EXPERIMENTAL RESULTS

This section illustrates the performance of the methodologies presented in Section III. Experiments have been conducted using a realistic data set from The Hamlyn Centre Laparoscopic [34]. The video sequence has a duration of 60.2 sec. in which the cardiac surface is affected by respiration and cardiac motion. Simulated experiments have been performed using a PC intel Core i7, 8GB RAM, and Nvidia GeForce GT 540M.

Firstly, it is worth noting that despite dealing with specular highlights it is not the topic of this work, this phenomenon has been kept in mind in order to avoid decreasing the available information. This process can be seen in Fig. 3, in which a step of detection using intensity information is carried out, after that, elimination of specular reflections are done, and finally, a inpainting method, using a priori knowledge, is applied in order to recover the loss information.

In order to evaluate both methodologies different factors have been taken into account: computational cost (average time per frame in seconds); convergence characteristic (average number of iterations per frame); average error which is given in mm.; and desirable characteristics (plausible deformation and behaviour with complex deformations). In our test, a grid of 6 x 6 control points is used with a defined separation space of 30 pixels.

In Fig. 6 part of the image sequence is shown, in which error decrease when regularization is used. Also, a well-defined displacement field is generated. In comparison with the
space of solution given as a result plausible deformation. In contrast, unplausible deformation is obtained when regularization term is not included.

![Fig. 4](image)

Fig. 4 Behaviour of both approaches during the image sequence. The \( \ell_1 \)- and \( \ell_2 \)- regularized versions allow restricting the space of solution given as a result plausible deformation. In contrast, unplausible deformation is obtained when regularization term is not included.

![Fig. 5](image)

Fig. 5 (Right to left) The zone of interest is handled a uniformly spaced lattice of \( 6 \times 6 \). A point of interest is selected as well as its four neighboring points which displacements in x and y coordinates are retrieved.

![Fig. 6](image)

Fig. 6 (a) Average time per frame using the regularized and unregularized version of both methodologies (b) Average number of iteration in the optimization process per Frame

TABLE I PERFORMANCE ANALYSIS

<table>
<thead>
<tr>
<th>Approach</th>
<th>Avg. Time per Frame [sec]</th>
<th>Avg. #Iterations per Frame</th>
<th>Avg. Error [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \ell_1 )-Unreg.</td>
<td>0.023</td>
<td>22</td>
<td>0.2098</td>
</tr>
<tr>
<td>( \ell_1 )-Reg.</td>
<td>0.011</td>
<td>15</td>
<td>0.0980</td>
</tr>
<tr>
<td>( \ell_2 )-Unreg.</td>
<td>0.031</td>
<td>24</td>
<td>0.2911</td>
</tr>
<tr>
<td>( \ell_2 )-Reg.</td>
<td>0.014</td>
<td>18</td>
<td>0.1002</td>
</tr>
</tbody>
</table>

V. DISCUSSION

This paper presented two approaches to compensate the physiological motion based on the class of: \( \ell_1 \)- and \( \ell_2 \)- regularized optimizations. In this context, the exploration of \( \ell_1 \)- was proposed as well as the benefits of using explicit regularization. According to the results, we suggest the use of the class of \( \ell_1 \)- regularized optimization for solving this problem, since it completed the tasks with the shortest time (0.011 sec.), the smallest error (0.0980 mm.) and demonstrated to be stable in presence of complex deformation. That is, this class offers more robustness and efficiency in comparison with the class of \( \ell_2 \). Although this work is focussed on the heart motion compensation, this methodology can also be applied in different application, in which deformation recovery is needed.
Some examples are suturing [35], tissue property estimation [36], virtual reality [37] among others. As future work, this methodology will be extended in 3 dimensions and unified with our work in force prediction [38] that avoid damaging the tissue.

ACKNOWLEDGMENT

“This work is supported by a FPU national scholarship from the Spanish Ministry of Education with reference AP2012-1943. Also, this work is part of the project DPI2011-29660-C04-01 and DPI2011-29660-C04-03 MINECO and with FEDER funds EC.

REFERENCES


