Sensorless Force Estimation using a Neuro-Vision-Based Approach for Robotic-Assisted Surgery

Angelica I. Aviles[1], Samar Alsaleh[2], Pilar Sobrevilla[3] and Alicia Casals[4]

Abstract— **This paper addresses the issue of lack of force feedback in robotic-assisted minimally invasive surgeries. Force is an important measure for surgeons in order to prevent intra-operative complications and tissue damage. Thus, an innovative neuro-vision based force estimation approach is proposed. Tissue surface displacement is first measured via minimization of an energy functional. A neuro approach is then used to establish a geometric-visual relation and estimate the applied force. The proposed approach eliminates the need of add-on sensors, carrying out biocompatibility studies and is applicable to tissues of any shape. Moreover, we provided an improvement from 15.14% to 56.16% over other approaches which demonstrate the potential of our proposal.**

I. INTRODUCTION

Robotic Assisted Minimally Invasive Surgery (RAMIS) systems, besides preserving the advantages of traditional laparoscopic procedures, offer additional benefits for both surgeons and patients. They enhance surgeons' precision by restoring hand-eye coordination and filtering physiologic tremors. Moreover, RAMIS is often associated with better patient outcomes and less pain, blood loss, scarring and recovery time [1]. However, these benefits have come at the expense of the surgeon's ability to feel, which eliminates all natural haptic feedback [2]. Force feedback is of utmost importance to prevent tissue damage, and increase both surgeon-patient transparency and procedure accuracy. Moreover, it is associated with task completion time and increased intraoperative injury [3].

The lack of force feedback has motivated some researchers to develop force sensors for restoring this feature to RAMIS [4], [5], [6]. However, biocompatibility and sterilizability problems, space restrictions, long-term stability, devices' high cost and the difficulty of adapting them to the surgical tool, put severe restrictions on their use in real clinical environments [7]. An alternative approach is to estimate the applied forces from the observable displacements of the deformable object. Since the camera integrated in the endoscope is able to infer tool-tissue interaction forces and prevent disturbances at the contact point, Vision-Based Force Measurement (VBFM) can be used.

The viability of using visual information for estimating forces has been proved in different scenarios. Authors in [8], [9] showed that vision-based force measurement provides a robust and non-contact approach for estimating force feedback, in macro- and micro-scale systems, in which displacement of the object's contour was enough to recover the applied forces. However, calculating the properties of the materials in these approaches is highly complex.

Force estimation in 3D space is more challenging and only few results have been reported. A 3D contact force estimator, based on the Extended Kalman filter (EKF) and a Lyapunov-based adaptation law, was proposed in [10]. Ammi et al. [11] estimated forces using a nonlinear mass-spring-damper model and snakes based visual tracking of cell deformations. Later on, a deformation-strain-stress relation was proposed in [12] to retrieve surface deformation using monocular images. However, as shown in [13], mass-spring models offer limited accuracy.

Most recently, researchers have used Neural Networks to improve force feedback estimation in different scenarios. Authors in [14] applied image processing techniques in conjunction with artificial neural networks to estimate the force applied to cells in 2D space. In a previous work [15], we proposed a 3D force estimation approach, in RAMIS scenarios, in which an information-force relation was created by means of a recurrent neural network. Additionally, the combination of 3D VBFM and artificial intelligence techniques was first reported in that work.

In this work, a neuro-vision based approach for estimating applied forces is presented in which tissue surface deformation is reconstructed by minimizing an energy funcional. To guarantee good performance even under occlusions, a geometric-visual relation is created using a neural approach. The proposed approach has the benefits of eliminating the need of using sensors, conducting biocompatibility studies and knowing tissue shape.

II. VISUAL-FORCE ESTIMATION STRATEGY

For tackling the problem of estimating the applied forces, a solution composed of two modules, illustrated in Fig. 1, is proposed. The former, vision-based approach, allows estimating the displacements produced during the deformation of the tissue. The latter, neuro approach, generates the information-force relation. In the remainder of this section, the stated modules are described in detail.

[1]* A.I. Aviles is with the Intelligent Robotics and Systems Group, Universitat Politècnica de Cataluya, 08034, Barcelona, Spain. angelica.ivone.aviles@upc.edu

[2] S. Alsaleh is with the Bioinformatics and Biomedical Computing Group, Department of Computer Science, George Washington University, 20052, Washington DC, U.S.A. sm57@gwu.edu

[3] P. Sobrevilla is with the Department of Applied Mathematics II, Universitat Politècnica de Cataluya, 08034, Barcelona, Spain. pilar.sobrevilla@upc.edu

[4] A. Casals is with the Institute for Bioengineering of Catalonia and with the Universitat Politècnica de Cataluya, 08028, Barcelona, Spain. alicia.casals@upc.edu
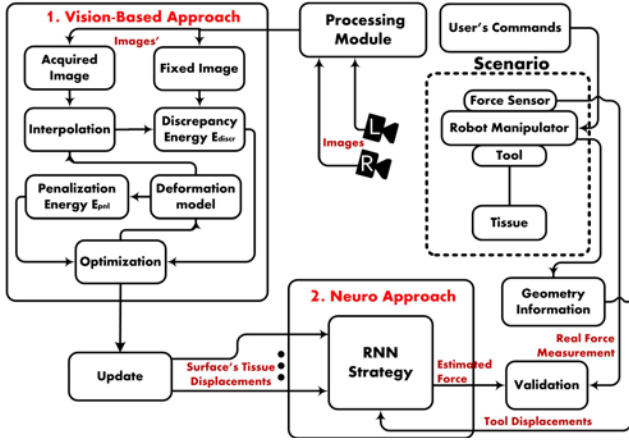
Fig. 1: Flow chart of the proposed solution for estimating the applied forces in RAMIS environments.
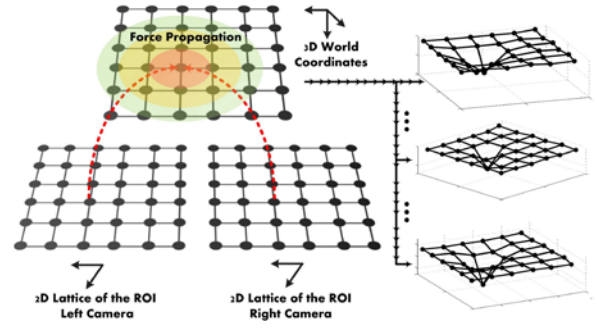

Fig. 2: 3D deformation structure is reconstructed, at each time instant, from the given lattices of corresponding pair of cameras (right side). Then, observed displacements can be seen due to the applied force (left side).

A. Vision-Based Approach

Consider a stereo-pair image sequence during the time, $t = 1, ..., N$, where definitions of the left and right images are given by $\mathbf{I}_l : \Omega_{\mathbf{I}_l} \to \mathsf{R}$ and $\mathbf{I}_r : \Omega_{\mathbf{I}_r} \to \mathsf{R}$, with compact support on Ω, respectively. Now, let $\mathbf{P}$ be a set of control points which belong to a uniform lattice on $\Omega_{\mathbf{I}_l}$ and $\Omega_{\mathbf{I}_r}$,
then, the three-dimensional grid is given by the projections of the corresponding points of each lattice (See left side Fig. 2). Thus, deformation structure is given by the repercussion of $\mathbf{P}$ (see right side Fig. 2). Mathematically, the changes produced on the tissue's surface are treated as an energy minimization problem, where the goal is to find $\mathbf{P}$ of

$$\min_{\mathbf{P}} \{ \mathbf{E}_{discr}(\mathbf{I}_l(\mathbf{d}_m(\mathbf{x};\mathbf{P}) + \mathbf{x}), \mathbf{I}_r(\mathbf{x})) + \gamma \mathbf{E}_{pnl}(\mathbf{d}_m(\mathbf{x};\mathbf{P})) \} \quad (1)$$

where E_{discr} measures the level of alignment between both images at each t, $\mathbf{E}_{pnl}$ represents the penalization term used to obtain a plausible transformation, γ gives a balance between $\mathbf{E}_{discr}$ and $\mathbf{E}_{pnl}$, and $\mathbf{x}$ is a vector in R^d, d-dimensional space. Moreover, for each $\mathbf{x} \in \mathsf{R}^d$ the disparity function, $\mathbf{d}_m$, is described by

$$\mathbf{d}_m(\mathbf{x};\mathbf{P}) = \sum_{j_1=0}^{n} ... \sum_{j_d=0}^{n} \mathbf{P}_{1,...j_d} \overset{contol\,points}{\prod_{k=1}^{d}} \underbrace{\xi_k(\mathbf{x}_k)}_{tensor\,product} \quad (2)$$

Where ξ_k are the B-splines basis functions of degree $n = 3$. These functions have been selected due to their ease of manipulation, low computational cost and optimal mathematical properties [16]. We now turn to reformulate the energy functional presented in Eq. 1. In this work, the sum of squared differences (SSD) and Tikhonov method are considered for evaluating the discrepancy and penalization terms, $\mathbf{E}_{discr}$ and $\mathbf{E}_{pnl}$, respectively. Thus, energy functional becomes

$$\overset{discrepancy}{\frac{1}{2}\sum_{\mathbf{x}\in\Omega}(\mathbf{I}_l(\mathbf{d}_m(\mathbf{x};\mathbf{P})) - \mathbf{I}_r(\mathbf{x}))^2} + \overset{penalization}{\gamma\sum_{i=1}^{d}\sum_{\mathbf{x}\in\Omega}|\nabla \mathbf{d}_{m_i}(\mathbf{x};\mathbf{P})|^2} \quad (3)$$

The set of control points minimized in Eq. 3 has been obtained by applying the Powells dogleg (DL) method because, besides combining the advantages of the Gauss-Newton and steepest descent methods, it offers better results than Levenberg-Marquardt with a fastest performance. Detailed description can be found in our previous works related to complex deformation [17], [18].

B. Neural Approach

In this paper, the applied forces estimation in RAMIS is carried out by using a Recurrent Neural Network (RNN). RNNs are a class of artificial neural networks (ANNs) that have feedback connections, which enable the networks to do temporal processing and learn sequences taking into account past information. Moreover, RNNs are capable of dealing with problems related to complex input-output relations as well as handling high noise-contaminated input data. Previous characteristics make RNNs very suitable for estimating forces in our system as they will allow us to create the visual-geometric-force relationship and use past information to improve final results.

The architecture used in this paper is illustrated in Fig. 3. Unlike our previous work [15], in which full feedback connections were considered, in this work, two main types of feedback, with delay, are introduced: global output-to-input feedback and local feedback in the input and output of the global loops. According to these feedback, three types of outputs can be defined: i) local delay unit $\mathbf{O^{LD}}$ ii) global delay unit $\mathbf{O^{GD}}$ ii) no-delay unit $\mathbf{O^{N-D}}$ of the $i-th$ neuron. They are expressed as follows

$$\mathbf{S}_{\mathbf{LD}_i}^{(t)} = \sum_j \mathbf{w}_{ij} + \mathbf{In}_j^{(t)} + \mathbf{b}_i + \sum_j \mathbf{w}_{ij}\mathbf{S}_{\mathbf{LD}_i}^{(t-1)}$$

$$\mathbf{S}_{\mathbf{GD}_i}^{(t)} = \mathbf{O^{LD}}^{(t)} + \sum \mathbf{S}_{\mathbf{GD}_i}^{(t-1)}$$

$$\mathbf{O^{LD}}^{(t)} = \varphi(\mathbf{S}_{\mathbf{LD}_i}^{(t)}) \quad \mathbf{O^{GD}}^{(t)} = \varphi(\mathbf{S}_{\mathbf{GD}_i}^{(t)})$$

$$\mathbf{O^{N-D}}^{(t)} = \varphi(\sum_j \mathbf{w}_{ij} + \mathbf{In}_j^{(t)} + \mathbf{b}_i) \quad (4)$$

where inputs are defined by $\mathbf{In}$, $\mathbf{b}$ denotes the bias, weights are represented by $\mathbf{w}$ and $\mathbf{S}$ refers to the current state.
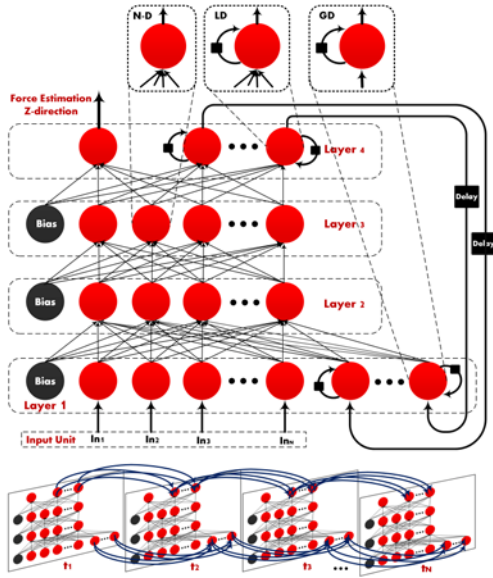
Fig. 3: Estimation of the applied forces is achieved by means of a RNN in which three types of output units can be identified (zoom in the upper row). Those units with delayed feedback save past information that helps to increase accuracy (bottom row).

Moreover, φ represents the activation function. As activation functions highly affect the performance of RNNs, two functions are used: identity activation function on the output layer and hyperbolic tangent sigmoid on the other layers to take advantage of the multilayer configuration. They were selected since in combination they improve the learning characteristics and increase accuracy and speed. One of the main concerns when using a RNN, because of the vanishing and the exploding gradient problems, is the training process. With the aim of obtaining faster convergence and stability, in this work the Levenberg-Marquardt method has been applied. Thus, the weight update is expressed by $\Delta \mathbf{w} = -(J^T J + \rho I)^{-1} J^T e$, where J and I are the Jacobian and identity matrices, ρ is a non-negative value that allows the Hessian be positive definite, and e is the vector error that is obtained between the output and the target.

III. EXPERIMENTAL RESULTS

The validation of the proposed approach has been carried out on a realistic scenario (upper left image of Fig. 5) composed of a pair of cameras from IDS, a Stäubli RX60B manipulator robot, and an ATI Gamma SI-32-2.5 force sensor. Furthermore, a fabric made of ECOFLEX 0030, which has mechanical properties similar to human tissue, has been used. The computational analysis of the proposed vision-based approach has shown a high performance with an average time per frame of 0.0060 seconds and a fast convergence (17 iterations in average). The top of Fig.5 shows a 3D plot of the observed displacement of the lattice vision output, while at the bottom of this figure, 2D graphs of displacements of the contact point (left) and its neighbors
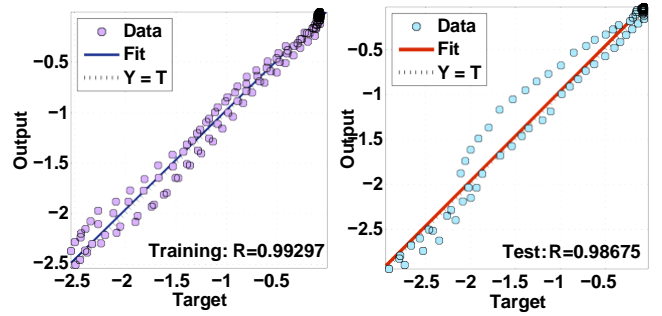

Fig. 4: Relationship between the RNN output and the targets obtained for the training and test datasets (left and right graphics, respectively).

(right) are displayed. This information is used as input of the proposed RNN. To evaluate the association strength between the real and estimated force measurements, a linear regression test was applied. To do so, 60% out of the total data was used for training and the remainder for test. In both cases, there were a very tight relationship between estimated and real force values, as shown by fitting of the points to the lines in Fig.4, and the R-values greater than and close to 0.99 for the training and test data, respectively. These results demonstrate the significant robustness and accuracy of the proposed RNN in estimating the applied forces

Additionally, root mean square (RMS) error was used to assess the performance of the proposed neuro-vision based approach, and compare it with the error obtained by other approaches. Red bar in left side of Fig. 6 identifies the RMS error obtained by the proposed approach, and is used as reference for comparison with the approaches reported in Table I. Moreover, blue line shows the percentage of improvement of our approach against these studies. As can be appreciated, the estimated forces closest to the target values are obtained with the system presented here, which provides an accuracy improvement ranging from 15.14% to 56.16% over other approaches and outperforms our previous proposal, [15], by a 5.26%. Apart from the improved force estimation accuracy (see right side Fig. 6) a significant reduction, in comparison to our previous work, of about 62.2% in computational cost was achieved. These results demonstrate the viability and significance of our proposed approach for RAMIS scenarios.

IV. CONCLUSION

We presented a novel approach to estimate the applied forces in the RAMIS context in which deformation of the tissue surface (given by the visual information) and geometric information (obtained by the robot manipulator) are used as input to the RNN. Moreover, the combination of two types of feedback allowed reducing both, the error, RMS= 0.0594, and the computational time. These factors, together with eliminating the need of conducting biocompatibility studies and knowing tissue shape, demonstrate the potential of the proposed approach. Moreover, we reported a significant improvement over the state of the art in force feedback from 15.14% to 56.16%.
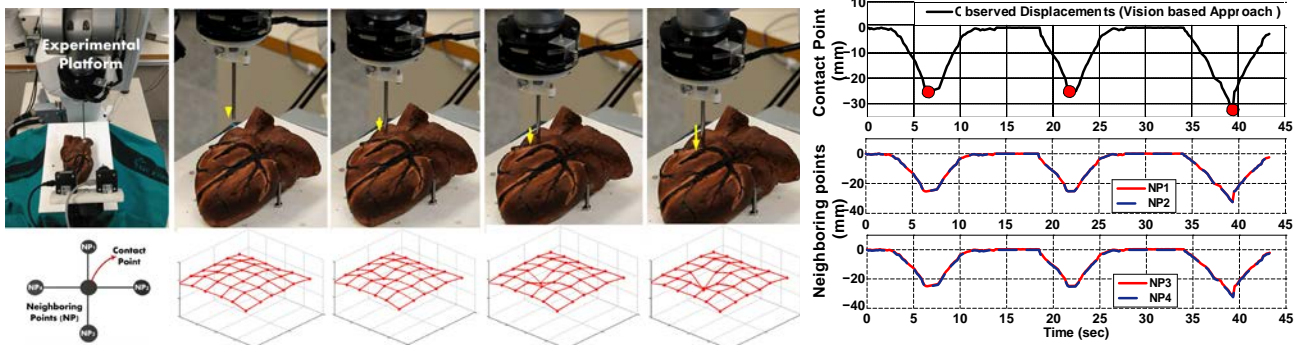
Fig. 5: At top of the first column, experimental platform can be seen, in which force is applied to the tissue at different time instants. 3D displacements, obtained from the vision based approach, are plotted (bottom of first column). Observed displacements of the contact point and its four neighboring points are also shown (second column).

TABLE I: Performance assessment, using the Root-Mean Square Error, of proposed and existing approaches.

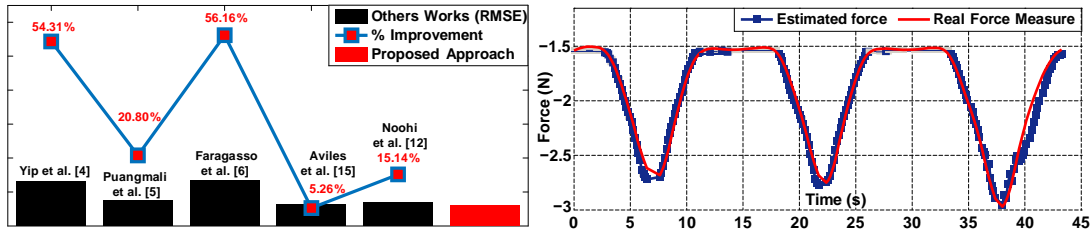| Compared Methods	Yip et al. [4]	Puangmali et al. [5]	Faragasso et al. [6]	Aviles et al. [15]	Noohi et al. [12]	**Proposed Approach**
RMS Error (N)	0.13	0.075	0.1355	0.0627	0.07	**0.0594**

Fig. 6: The RMS Error reported by all approaches in Table I. Red bar identifies RMS Error obtained by the proposed approach and is used as reference for the graph (at left side). Percentage of change against other studies is also shown. Comparison between the real and estimated forces, in Z-direction, are shown (at right side).

REFERENCES

[1] E.B. Wilson, H. Bagshahi, and V.D. Woodruff, Overview of General Advantages, Limitations, and Strategies *Book Chapter, Robotics in General Surgery*, Springer, 2014.

[2] A.M. Okamura, Haptic feedback in robot-assisted minimally invasive surgery, *Current Opinion in Urology, 19(1):102-107*, 2009.

[3] C. Pacchierotti, A. Tirmizi and D. Prattichizzo, Improving transparency in teleoperation by means of cutaneous tactile force feedback, *ACM Transactions on Applied Perception*, 2014.

[4] M.C. Yip, S.G. Yuen and R.D. Howe, A Robust Uniaxial Force Sensor for Minimally Invasive Surgery, *IEEE Trans. on Biomedical Engineering*, pp. 1008-1011, 2010.

[5] P. Puangmali, H. Liu, L.D. Seneviratne, P. Dasgupta and K. Althoefer, Miniature 3-Axis Distal Force Sensor for Minimally Invasive Surgical Palpation, *IEEE Trans. on Mechatronics*, pp. 646-656, 2012.

[6] A. Faragasso, J. Bimbo, Y. Noh, A. Jiang, S. Sareh, H. Liu, H.A. Wurdemann and K. Althoefer, Novel Uniaxial Force Sensor based on Visual Information for Minimally Invasive Surgery, *IEEE International Conference on Robotics and Automation*, 2014.

[7] S. Sokhanvar, J. Dargahi, S. Najarian and S. Arbatani, Clinical and Regulatory Challenges for Medical Devices *Tactile Sensing and Displays. Haptic Feedback for Minimally Invasive Surgery and Robotics*, Wiley Publications, 2012.

[8] M.A. Greminger and B.J. Nelson, Vision-based force measurement *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2004.

[9] J. Kim, F. Janabi-Sharifi and J. Kim, A Haptic Interaction Method Using Visual Information and Physically Based Modeling *IEEE Transactions on Mechatronics, Vol. 15, No. 4*, 2010.

[10] J. Jung, J. Lee and K. Huh, Robust contact force estimation for robot manipulators in three-dimensional space *Journal of Mechanical Engineering Science, vol. 220*, 2006.

[11] M. Ammi, H. Ladjal and A. Ferreira, Evaluation of 3D Pseudo-Haptic Rendering using Vision for Cell Micromanipulation, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006.

[12] E. Noohi, S. Parastegari and M. Zefran, Using Monocular Images to Estimate Interaction Forces During Minimally Invasive Surgery, *IEEE International Conference on Intelligent Robots and Systems*, 2014.

[13] A. E. Kerdok, S. M. Cotin, M. P. Ottensmeyer, A. M. Galea, R. D. Howe, and S. L. Dawson, Truth cube: Establishing physical standards for soft tissue simulation, *Medical Image Analysis, Volume 7*, 2003.

[14] F. Karimirad, S. Chauhan and B. Shirinzadeh, Vision-based force measurement using neural networks for biological cell microinjection *Journal of Biomechanics, Volume 47, Issue 5, pp. 1157-1163*, 2014.

[15] A.I. Aviles, A. Marban, P. Sobrevilla, J. Fernandez and A. Casals, A Recurrent Neural Network Approach for 3D Vision-Based Force Estimation, *IEEE International Conference on Image Processing Theory, Tools and Applications* , 2014.

[16] A.I. Aviles and A. Casals. Interpolation Based Deformation Model for Minimally Invasive Beating Heart Surgery. *In Book Chapter IFMBE Proceedings*, Vol.41, Springer International Publishing, 2013.

[17] A.I. Aviles, P. Sobrevilla and A. Casals, An Approach for Physiological Motion Compensation in Robotic-Assisted Cardiac Surgery, *Journal in Experimental and Clinical Cardiology*, 2014.

[18] A.I. Aviles, P. Sobrevilla and A. Casals, Unconstrained .el−Regularized Minimization with Interpolated Transformation for Heart Motion Compensation, *International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014.