



HAL
open science

Digital twins that learn and correct themselves

Beatriz Moya, Alberto Badías, Icíar Alfaro, Francisco Chinesta, Elías Cueto

► **To cite this version:**

Beatriz Moya, Alberto Badías, Icíar Alfaro, Francisco Chinesta, Elías Cueto. Digital twins that learn and correct themselves. *International Journal for Numerical Methods in Engineering*, 2022, 123 (13), pp.3034-3044. 10.1002/nme.6535 . hal-03709235

HAL Id: hal-03709235

<https://hal.science/hal-03709235>

Submitted on 29 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Digital twins that learn and correct themselves

Beatriz Moya¹ | Alberto Badiás¹ | Icíar Alfaro¹ | Francisco Chinesta² | Elías Cueto¹ 

¹Aragon Institute of Engineering Research, Universidad de Zaragoza, Zaragoza, Spain

²ESI Group Chair, PIMM Lab - Procédés et Ingénierie en Mécanique et Matériaux, Arts et Metiers Institute of Technology, Paris, France

Correspondence

Elías Cueto, Aragon Institute of Engineering Research, Universidad de Zaragoza, Edificio Betancourt, Maria de Luna, s.n. 50018 Zaragoza, Spain.
Email: ecueto@unizar.es

Funding information

Ministerio de Economía y Competitividad, Grant/Award Number: CICYT-DPI2017-85139-C2-1-R

Abstract

Digital twins can be defined as digital representations of physical entities that employ real-time data to enable understanding of the operating conditions of these entities. Here we present a particular type of digital twin that involves a combination of computer vision, scientific machine learning, and augmented reality. This novel digital twin is able, therefore, to see, to interpret what it sees—and, if necessary, to correct the model it is equipped with—and presents the resulting information in the form of augmented reality. The computer vision capabilities allow the twin to receive data continuously. As any other digital twin, it is equipped with one or more models so as to assimilate data. However, if persistent deviations from the predicted values are found, the proposed methodology is able to correct on the fly the existing models, so as to accommodate them to the measured reality. Finally, the suggested methodology is completed with augmented reality capabilities so as to render a completely new type of digital twin. These concepts are tested against a proof-of-concept model consisting on a nonlinear, hyperelastic beam subjected to moving loads whose exact position is to be determined.

KEYWORDS

augmented reality, computer vision, digital twins, scientific machine learning

1 | INTRODUCTION

The concept of digital twin is rooted in the concept of (hardware) twin established by NASA in the Apollo mission, where a twin of the spaceship was kept in Earth for security purposes. Its digital counterpart appears to be first cited again in a NASA report.¹ It has now become ubiquitous in industry, particularly with the burgeoning presence of the Internet of Things.² Essentially, it refers to a class of digital model which is connected to its physical twin counterpart by means of sensors, that produce real-time data. These data are expected to be assimilated by the digital model so as to allow to obtain knowledge about the physical system's functioning.

The concept is of course closely related to that of dynamic data-driven application systems (DDDAS).³⁻⁵ Darema, who coined the term,^{6,7} highlighted the symbiotic collaboration between data and simulation. The interested reader could consult the web page <http://dddas.org>, where a very instructive discussion about DDDAS and related concepts (such as cyber-physical systems,⁸ for instance) is made. In essence, what distinguishes a DDDAS from other types of digital twins is its ability not just to assimilate data into an existing application, but also to allow these data to steer the experimental campaign.

For complex systems, the need for real-time simulation has attracted the attention of the reduced order modeling community. Indeed, for very fast feedback responses, nowadays computing capabilities could not attain these extremely fast rates—that depend obviously on the particular application. Thus, there is a vast corps of literature on the use of reduced order models to overcome these difficulties.⁹⁻¹⁴

Here we explore a new capability for the design of efficient digital twins. Very often, these twins produce inaccurate predictions, whose systematic deviation from measurements cannot be attributed to the presence of noise in the experimental devices. Systems evolve during their service life, so that models employed for data assimilation are no longer valid and must be corrected. If this correction can be accomplished by the digital twin itself we speak of a hybrid twin.¹⁵

The so-called *fourth paradigm of science*¹⁶ refers to the capacity of searching through huge databases and extract valid information by removing correlations on it. Here, on the contrary, we study systems able to go one step further and, by means of scientific machine learning¹⁷ be able to unveil models from the available data. Sometimes understood restrictively as the employ of machine learning to solve PDEs,¹⁸ this discipline has been defined as *a core part of Artificial Intelligence and a computational technology that can be trained, with scientific data, to augment or automate human skills*.¹⁷ It constitutes, therefore, the cornerstone of our approach to hybrid twins here developed. So to speak, the type of digital twins we advocate for can be considered as a sort of *intelligence augmentation* device.¹⁹

Intelligence augmentation (IA, do not confuse with artificial intelligence, AI), can be defined as the set of *technologies that enhance human productivity and improve or restore capabilities of the human body or mind*.²⁰ The concept was first set up by Douglas Engelbart back to 1962.²¹ Essentially, our approach to the digital twin concept has many features in common to this concept of intelligence augmentation. Indeed, Augmented Reality can be seen as a cognitive technology, in the sense that it constitutes the ultimate form of computer interface, allowing for a seamlessly way of transmitting information to the user. This information, in this case, refers to the physics taking place in the observed phenomenon.

Therefore, our approach to digital twins comprises systems able to see—through computer vision technologies, to interpret what they see—through scientific machine learning technologies, and to seamlessly transmit to the user the essential hidden information about these physics. This transmission is made through AR technologies. Previous attempts at incorporating machine learning to the determination of constitutive laws can be found, among others, in the work of Michopoulos and collaborators.^{22,23}

In what follows we will describe the essential ingredients of this approach. First, an overview, for completeness, of the basics of computer vision, that will constitute our data acquisition system. Then, the scientific machine learning approach here employed. This is not, of course, the only approach that can be employed. Artificial neural networks (ANN), for instance, could replace this core capability of our system, although their unpredictability under some circumstances is well known. Finally, we will describe the AR implementation of the results.

The structure of the article is as follows. In Section 2 we describe the proof-of-concept problem that we have implemented. Basically, it is a foam beam being loaded at variable positions. The system is expected to located the exact position of this load, even if it is not visible from the video stream. In Section 3 we describe briefly the concept of hybrid twin, that is, the one that is able to correct itself by leveraging data. In Section 4 we describe our approach to the problem of correcting the model governing the assimilation procedure. Section 5 describes the mathematical problem of finding the position of the load as a particular instance of data assimilation. Finally, in Section 6 we analyze the performance of the method. The article is closed by the usual conclusions section.

2 | PROOF OF CONCEPT

Our proof-of-concept system can be thought of as a structural monitoring system in which a simply supported beam is subjected to moving loads, see Figure 1. The dimensions of the beam are $797.5 \times 200 \times 105 \text{ mm}^3$.

The beam was first employed in one of the authors' previous works.²⁴ During the experimental campaign, its constitutive behavior was found to be close to a Saint-Venant Kirchhoff law, with Young's modulus $E = 0.11 \text{ MPa}$ and Poisson's coefficient $\nu = 0.2$. Despite the well-known limitations of this model, for the load levels employed in the experiments, this model offered much better results than neo-Hookean or Mooney-Rivlin models. Mild viscoelastic effects were nevertheless found that do not greatly affect the results if the time scale at which we translate the load is shorter than the typical relaxation time of the material.



FIGURE 1 A view of the foam beam at the lab. It consists of a simply supported beam whose constitutive behavior is known to fit well to a Kirchhoff–Saint Venant law. The thick black lines help to track selected points in the image

The beam was meshed with 18,938 nodes and 98,863 linear tetrahedral elements. Supports were modeled as pinned, allowing the free rotation of the beam, but not the horizontal displacement. Loads of 15 N were applied with the help of lead disks with 80-mm diameter. These loads were nevertheless considered as punctual in the simulations.

The goal of the constructed system was to predict the exact position of the load, together with an AR plotting of the displacement, strain, or stress fields on top of the beam’s geometry. As a restrictive assumption, this should be done under real time constraints. This means that we must accomplish to the number of frames per second at which the video is recorder (usually 30–60 fps).

Under this rationale, the displacement field of the beam is assumed to be of the form

$$\mathbf{u} = \mathbf{u}(\mathbf{X}, Q),$$

where $\mathbf{X} \in \Omega \subset \mathbb{R}^3$ represents the position of a given point in the undeformed frame of reference, and $Q \in \Gamma_t \subset \mathbb{R}^3$ represents the particular position of the applied load. In general, we assume a simplified description where the position of the load is always well centered on the upper surface of the beam so that

$$\mathbf{u} = \mathbf{u}(\mathbf{X}, Q),$$

with $Q \in \mathbb{R}$ a one-dimensional coordinate along the plane of symmetry of the upper surface of the beam. This parametric form of the displacement field is precisely at the origin of the hybrid twin concept, that we describe next.

3 | THE CONCEPT OF HYBRID TWIN

In general, a digital twin includes a model describing the physics taking place. Assume that this model takes a particular form

$$\mathbf{u}(\mathbf{X}, t; \boldsymbol{\mu}) = \mathbf{A}(\mathbf{X}, t; \boldsymbol{\mu}), \quad (1)$$

where $\boldsymbol{\mu} \in \mathcal{D} \subset \mathbb{R}^{n_{\text{param}}}$ represents a set of parameters determining its behavior and t is time. In our case, for the sake of simplicity, we assume that $\boldsymbol{\mu}$ is composed solely by $Q \in [a, b] \subset \mathbb{R}$. For the problem under consideration, the model \mathbf{A} (here assumed linear, but that could also depend on the solution itself when the model is nonlinear) could be, for instance, the well-known Euler–Bernoulli–Navier (EBN) model for beam bending. This model is well known to describe accurately the bending of long beams under small strain assumptions. Given the parametric nature of the model, an approximation constructed by any model order reduction technique could be envisaged.²⁵⁻³⁰ In this case, however, given the simplicity of the Euler–Bernoulli–Navier model, we will take it as a starting point in our developments.

Given that it is well known that this beam obeys a nonlinear description of the movement, the assumed model will soon reveal its lack of accuracy, as will be reported in Section 6. If this is the case, given some unfitted measurements, the model will need to be corrected, by assuming

$$\mathbf{u}(\mathbf{X}, t; \boldsymbol{\mu}) = \mathbf{A}(\mathbf{X}, t; \boldsymbol{\mu}) + \mathbf{B}(\mathbf{X}, t; \boldsymbol{\mu}). \quad (2)$$

In this case, we assume that the correction term $\mathbf{B}(\mathbf{X}, t; \boldsymbol{\mu})$ also depends on the set of parameters $\boldsymbol{\mu}$. It is worth noting that the original form of the Hybrid Twin¹⁵ does not consider this dependence. This does not alter, however, the spirit of the concept.

Still an additional term could be considered in the structure of a hybrid twin. In general, every measurement will incorporate noise:

$$\mathbf{u}(\mathbf{X}, t; \boldsymbol{\mu}) = \mathbf{A}(\mathbf{X}, t; \boldsymbol{\mu}) + \mathbf{B}(\mathbf{X}, t; \boldsymbol{\mu}) + \mathbf{R}(t). \quad (3)$$

This noise could come as a consequence of experimental measurement errors or as a consequence of external actions over which we have no control. In general, there is a long tradition in engineering sciences on how to filter this type of noise, through the use of more or less sophisticated filters. Examples exist also for filtering procedures constructed over reduced order models.³¹

There is, however, an additional possible source of noise. It is well known that model order reduction leads, in general, to the generation of non-Markovian, history-dependent terms in the resulting model, and also noise. This noise comes from the fact of ignoring details from the unresolved variables. This is the well-known Mori–Zwanzig theory.^{32,33} When learning physics from data, these terms could result in important deviation from the expected behavior of the system.³⁴ For the problem at hand, however, the assumed theory is considered as a microscopic enough description for which no additional detail is necessary, such as, for instance, microstructural representative volume elements, and so on.

Still an additional term could be considered on the most general form of a hybrid twin: a control term $\mathbf{C}(t)$,

$$\mathbf{u}(\mathbf{X}, t; \boldsymbol{\mu}) = \mathbf{A}(\mathbf{X}, t; \boldsymbol{\mu}) + \mathbf{B}(\mathbf{X}, t; \boldsymbol{\mu}) + \mathbf{C}(t) + \mathbf{R}(t). \quad (4)$$

For this particular case, however, in which we consider a quasi-static behavior, no control term has been included.

The term “hybrid” comes precisely from the fact that in Equation (4) the response of the system comes mainly from two sources: a theoretical, even phenomenological model, encompassed in the contribution \mathbf{A} , and a data-driven correction \mathbf{B} . This is precisely the main difference with respect to classical digital twins. The procedure to correct the model is studied in the following section.

4 | THE CORRECTION PROCEDURE

The correction procedure is based, on one hand, on experimental measurements coming from video recordings of the loading process. On the other hand, a reconstruction of the high dimensional displacement field given by Equation (2) is needed. This parametric expression needs to be sampled experimentally, and this may lead to the well-known curse of dimensionality—an exponential increase of the number of samples with the number of dimensions of the phase space. To avoid this curse, we employed a technique coined as *sparse-proper* generalized decomposition, *s*-PGD.³⁵ This technique allows for a high-dimensional reconstruction of the sought field \mathbf{u} with a minimum number of sampling points and, notably, it involves only a sequence of three dimensional and one-dimensional problems (for the spatial part of the problem and the parametric dependency, respectively). Thus, there is no need to solve problems defined in the phase space of the model, 4D in this case.

4.1 | The vision system of the twin

To allow the twin to see, we employ a stereo camera *Zed Mini* from *Stereo Labs* (<https://www.stereolabs.com/zed-mini/>). This type of camera provides us with an estimation of the depth from the objective for each object in the scene. A sketch of the functioning of such a camera is shown in Figure 3. The interested reader is referred to some of our previous works and the references therein for a detailed explanation on how the camera works.³⁶

Essentially, the advantage of a stereo camera comes from the fact that a triangulation of the scene is possible at each frame, thus providing the depth field at each frame. In the absence of such a facility, a standard camera could equally

been used.³⁷ In general, a point \mathbf{P} of the scene is captured by both cameras giving a projection \mathbf{p}_{left} and $\mathbf{p}_{\text{right}}$, respectively. For a *frontal parallel* camera both points will thus appear on the same vertical position, namely $y_{\mathbf{p}_{\text{left}}} = y_{\mathbf{p}_{\text{right}}}$. The difference between the horizontal coordinates defines the disparity $d = x_{\mathbf{p}_{\text{left}}} - x_{\mathbf{p}_{\text{right}}}$, which can be related to depth by the equation

$$Z = \frac{fT}{x_{\mathbf{p}_{\text{left}}} - x_{\mathbf{p}_{\text{right}}}},$$

where f is the focal length (assuming that both cameras are identical and have, therefore, the same value) and T is the displacement between camera centers of projection.

Once a three-dimensional reconstruction of the scene is made, the next step is to detect and track a minimal number of points in the beam. These are referred to as *features* in the computer vision community. While there exists a vast corps of literature on how to accomplish this task, here we have employed the OpenCV library³⁸ and, in particular, the feature detecting algorithm by Shi and Tomasi.³⁹ As will be noticed, this algorithm is able to track the corners in the black pattern drawn in the beam.

4.2 | Sparse proper generalized decomposition

Once a predefined set of points or features have been measured, it is time to obtain the correction for the model $\mathbf{A}(\mathbf{X}, t; \boldsymbol{\mu})$, $\mathbf{B}(\mathbf{X}, t; \boldsymbol{\mu})$, see Equation (2). This correction needs not to be performed at video speeds. It could be executed once systematic, biased deviations from the predictions given by the model are detected.

Our unknown objective function will therefore be the sought correction,

$$\mathbf{B}(\mathbf{X}, t; \boldsymbol{\mu}) = \mathbf{u}(\mathbf{X}, t; \boldsymbol{\mu}) - \mathbf{A}(\mathbf{X}, t; \boldsymbol{\mu}).$$

At a particular time instant t we will look therefore for an approximation to this correction, $\hat{\mathbf{B}}$, such that it minimizes

$$\int_{\mathcal{D}} \int_{\Omega} \omega^*(\mathbf{X}; \boldsymbol{\mu}) (\hat{\mathbf{B}}(\mathbf{X}; \boldsymbol{\mu}) - \mathbf{B}(\mathbf{X}, t; \boldsymbol{\mu})) d\Omega d\boldsymbol{\mu},$$

where $w^*(\mathbf{X}; \boldsymbol{\mu}) \in C^0(\Omega, \mathcal{D})$ is an arbitrary test function.

The main ingredient of the proper generalized decomposition to avoid the curse of dimensionality is to assume the sought field in a separate form,

$$\hat{\mathbf{B}}(\mathbf{X}; \boldsymbol{\mu}) \approx \sum_{i=1}^{n_{\text{modes}}} \mathbf{F}_i(\mathbf{X}) \cdot G_i^1(\mu_1) \cdot \dots \cdot G_i^{n_{\text{param}}}(\mu_{n_{\text{param}}}),$$

where $\mathbf{F}_i(\mathbf{X})$ and $G_i^j(\mu_j)$, $j = 1, \dots, n_{\text{param}}$ are sought functions, a priori unknown, coined as *modes* due to their similitude with proper orthogonal decomposition eigenvectors.

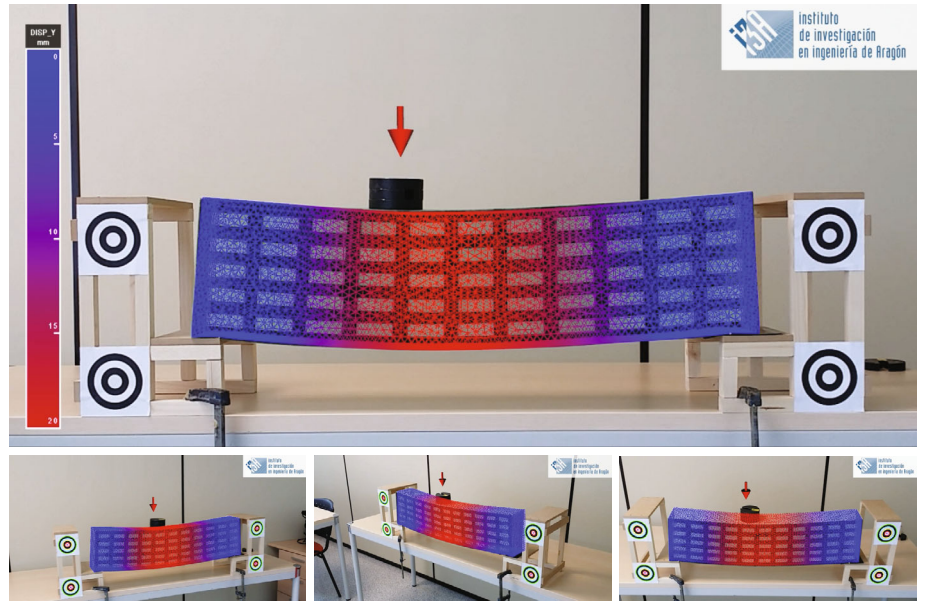
It is well known that this assumption of affinity of the sought approximation, is not appropriate for every system. Those in which the number of terms n_{modes} is huge are known as *nonseparable*. They are therefore, harder to approximate. To overcome this issue, improved PGD approximations can be employed, such as local-PGD approximations, for instance.^{40,41}

These functions are determined by first projecting them onto appropriate finite element spaces. Then, a greedy algorithm is employed so as to compute each of these sums,

$$\hat{\mathbf{B}}^m(\mathbf{X}; \boldsymbol{\mu}) = \hat{\mathbf{B}}^{m-1}(\mathbf{X}; \boldsymbol{\mu}) + \mathbf{F}_m(\mathbf{X}) \cdot G_m^1(\mu_1) \cdot \dots \cdot G_m^{n_{\text{param}}}(\mu_{n_{\text{param}}}). \quad (5)$$

In general, our stopping criterion is to enrich until the modulus of the enrichment at iteration m is lower than 1% of the modulus of the first mode. All of our experiments converged in less than 10 modes.

FIGURE 2 Results of the analysis of the foam beam. The system should be able to determine the position of the load—indicated as a red arrow in the images—while depicting the desired field on top of the beam’s geometry. Here, the legend represents the vertical component of the displacement, u_y . The system should be robust respect any displacement of the camera and even partial or general occlusions of the beam. The black circles are fiducial markers employed to calibrate the system, but they are not necessary during the operating regime of the system



For each sum, each function is computed iteratively by resorting to a fixed-point algorithm. In order to obtain a symmetric expression, it is a common practice to choose

$$\begin{aligned}
 w^*(\mathbf{X}; \boldsymbol{\mu}) = \hat{\mathbf{B}}^*(\mathbf{X}; \boldsymbol{\mu}) = & \mathbf{F}^*(\mathbf{X}) \cdot G_m^1(\mu_1) \cdot \dots \cdot G_m^{n_{\text{param}}}(\mu_{n_{\text{param}}}) \\
 & + \mathbf{F}_m(\mathbf{X}) \cdot G_m^1(\mu_1) \cdot \dots \cdot G_m^{n_{\text{param}}}(\mu_{n_{\text{param}}}) + \dots \\
 & + \mathbf{F}_m(\mathbf{X}) \cdot G_m^1(\mu_1) \cdot \dots \cdot G_m^{n_{\text{param}}}(\mu_{n_{\text{param}}}).
 \end{aligned} \tag{6}$$

This approach is equivalent to POD⁴² in two dimensions and for elliptic operators. There is no formal proof, however, on its ability to obtain parsimonious models (i.e., with a minimal number of terms) in dimensions higher than two.

In our approach, since we have a finite—usually small—set of measurements, it seems reasonable to employ Dirac delta functions as test functions, thus arriving at a collocation approach:

$$w^*(\mathbf{X}; \boldsymbol{\mu}) = \hat{\mathbf{B}}^*(\mathbf{X}; \boldsymbol{\mu}) \sum_{j=1}^{n_{\text{meas}}} \delta(\mathbf{X}_j; \boldsymbol{\mu}_j).$$

The matrix expression of the resulting formulation can be consulted in Ibañez et al.³⁵

Special attention needs to be paid to the chosen approximation spaces.³⁵ Here, we have employed standard finite elements in space, while the parameter space $Q \in [0, 795]$ mm is approximated by means of global Kriging procedures. Since the tracked points belong to the front, visible side of the beam, a simple grid of 13×5 nodes has been defined on the rectangular side of the beam. Each element is thus 66.25×50 mm long, which is enough for the level of accuracy here pursued. Of course, the main limitation of the technique arises at the low data limit, where a very limited number of experimental values are available. If needed, this grid can be refined without any limitation.

5 | MATHEMATICAL FORMULATION OF THE DATA ASSIMILATION PROBLEM

Out of the correction loop just described, which is invoked automatically every time that a biased deviation from the predictions is detected, the digital twin is expected to be operating normally, providing the user with information regarding the solution field (displacement or its derivatives, strain or stresses) and the location of the load, as seen in Figure 2.

This problem is formulated as a data assimilation problem, in the following manner. From the camera recording, sketched in Figure 3, a set of displacements is measured. These could be the same employed in the previous section

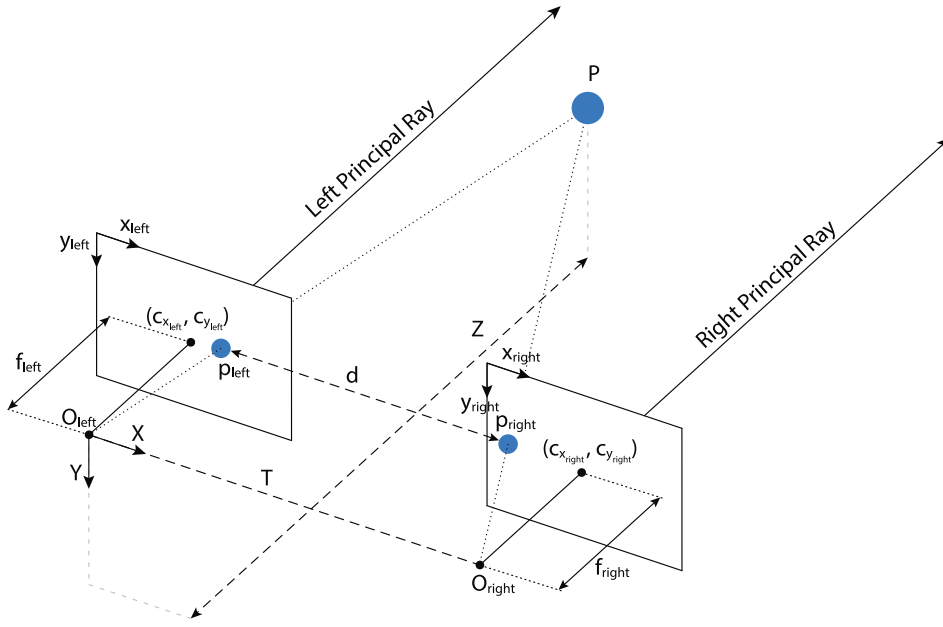


FIGURE 3 Schematics of a stereo camera system. The triangulation process is assumed to be coplanar. f_{left} is the left focal distance, $c_{x_{\text{left}}}, c_{y_{\text{left}}}$ are the points where the left principal ray is intersecting the left image plane, O_{left} is the origin of the left principal ray, T is the displacement between left and right camera centers and d is the disparity (difference) of the horizontal coordinates of the left and right projections of point P

for model correction or not. These lectures will be denoted as $\mathbf{u}^*(\boldsymbol{\mu})$.²⁴ In this framework, the continuous \mathbf{u} problem—the corrected Euler–Bernoulli–Navier beam problem—can, after discretization, be cast into the following two discretized PDEs,

$$\begin{aligned} \text{State equation: } & K(\mathbf{u}, \boldsymbol{\mu})\mathbf{u}(\mathbf{x}, \boldsymbol{\mu}) = \mathbf{f}(\boldsymbol{\mu}), \\ \text{Observation equation: } & \mathbf{u}^{\text{obs}}(\boldsymbol{\mu}) = C(\boldsymbol{\mu})\mathbf{u}(\mathbf{x}, \boldsymbol{\mu}). \end{aligned} \quad (7)$$

It is worth noting that the stiffness matrix $K(\mathbf{u}, \boldsymbol{\mu})$ could depend (possibly in a nonlinear manner, especially after correction) on the unknown field \mathbf{u} but also on the parameters $\boldsymbol{\mu}$. $\mathbf{u} \in \mathbb{R}^{n_{\text{FOM}}}$ represents the set of nodal (finite element full-order model) degrees of freedom and $\mathbf{u}^{\text{obs}} \in \mathbb{R}^{n_{\text{obs}}}$ represents the set of observations taken from the video stream. The observation matrix C is assumed to be linear and is usually simply Boolean and often $\boldsymbol{\mu}$ -independent.

We thus formulate the problem as a deterministic one, in the form

$$\boldsymbol{\mu}^* = \arg \min_{\boldsymbol{\mu} \in \mathbb{R}^{n_{\text{param}}}} \frac{1}{2} \|\mathbf{u}^{\text{obs}} - C(\boldsymbol{\mu})\mathbf{u}(\mathbf{x}, \boldsymbol{\mu})\|_2^2, \quad (8)$$

that is, we determine the value of the parameters that minimize the discrepancy between theoretical predictions of the (possibly corrected) model, $\mathbf{u}(\boldsymbol{\mu})$, and experimental measurements, \mathbf{u}^{obs} at particular locations given by C . Tikhonov regularization methods could be necessary in some circumstances.⁴³ In the experiments analyzed during this work, however, no need of any regularization was found.

Given the severe real-time restrictions (feedback rates between 30-60 Hz), instead of a closed-form model, like in this example, a reduced-order approximation to \mathbf{u} can be advantageously constructed by PGD. Equation (8) is then solved by employing the Levenberg–Marquardt method.^{9,44,45} The sensibilities of the solution with respect to the parameters can be found thanks to the separated structure of PGD approximations of both model and correction as

$$\frac{\partial \mathbf{u}}{\partial \mu_j}(\mathbf{X}, \boldsymbol{\mu}) \approx \sum_{i=1}^n \mathbf{F}_i(\mathbf{X}) \cdot G_i^1(\mu_1) \cdot \dots \cdot \frac{\partial G_i^j(\mu_j)}{\partial \mu_j} \cdot \dots \cdot G_i^{n_{\text{param}}}(\mu_{n_{\text{param}}}).$$

These sensibilities could be computed off-line and stored in memory. Only during runtime they are evaluated.

6 | INITIAL ANALYSIS OF THE PERFORMANCE

As mentioned before, the model \mathbf{A} is taken in this case, for simplicity, as an Euler–Bernoulli–Navier model, which is exactly separable. Corrections to this model are computed online. Modes \mathbf{F}_m and G_m in Equation (5), $m = 1, \dots, 9$ are

FIGURE 4 One-dimensional modes in the load position coordinate $G_m(Q)$. The first nine modes are represented

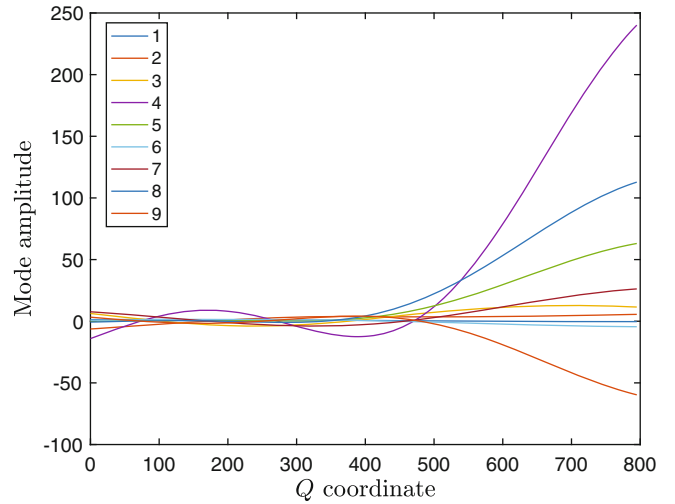
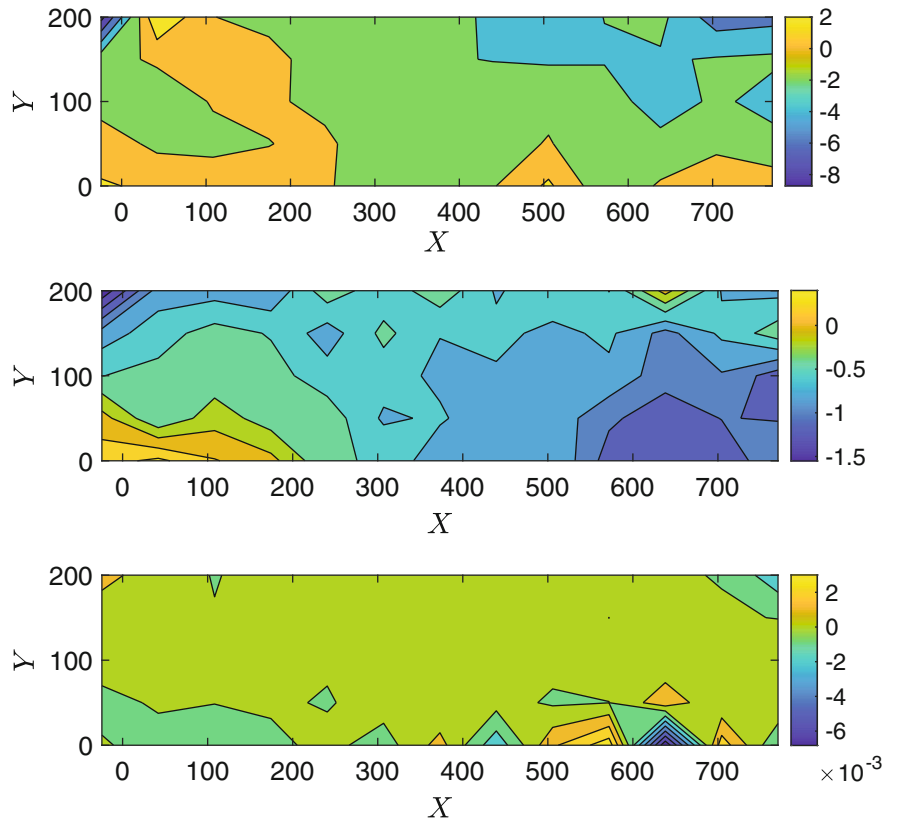


FIGURE 5 Two-dimensional modes in the physical coordinates $F_m(\mathbf{X})$. Modes $m = 1, 2,$ and 9 are represented



shown in Figures 4 and 5. The stopping criterion, as mentioned before, is the reach of a mode whose modulus is smaller than 1% of that of the first mode.

The typical image obtained by the computer vision acquisition system is shown in Figure 6. White points in the corners of the black grid denote that a particular corner is being tracked and its displacement measured. Note how some of them (bottom right part of the beam) do not appear in some frames of the video. This does not affect to the robustness of the method. Even partial occlusions during short time lapses are supported by the method without problem.

Figure 7 represents graphically the experimental measurements, predictions done by the EBN model and, finally, their correction by sparse-PGD techniques. Note the good agreement between circles and crosses, that is, between experimental measurements and the corrected predictions made by the digital twin. In this case, 10 modes are enough to correct the poor prediction done by the EBN model.

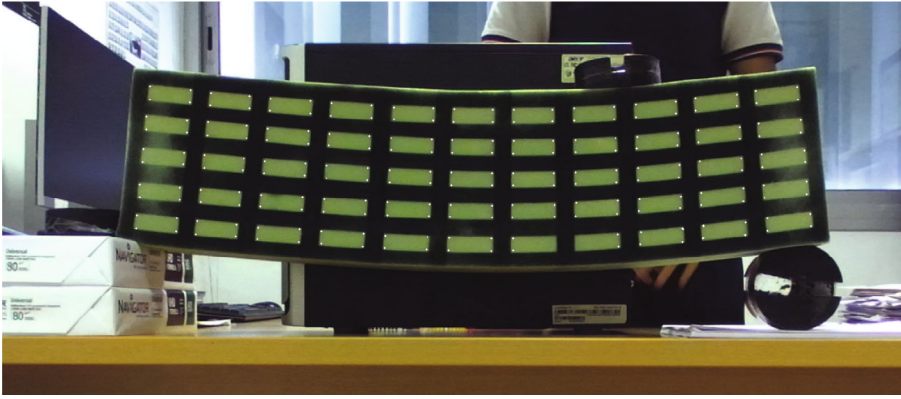


FIGURE 6 Frame of the video stream in which we observe the tracking of the corners in the black grid, represented as white points (these do not exist in the physical beam)

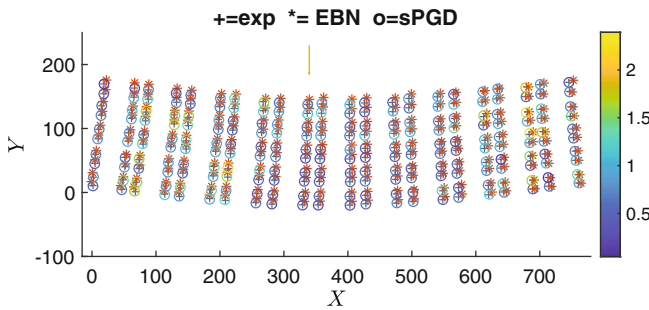


FIGURE 7 Predicted and measured displacement. Crosses represent the experimental measurements at the black grid corner points, represented in white in Figure 6. The displacement prediction made by the Euler–Bernoulli–Navier model is represented by asterisks. Finally, after the computation of a correction to the EBN model, the resulting prediction is represented by circles. The legend represents the error value, in millimeters

Video was recorded at a resolution of 1920×1080 pixels. The experimental error is therefore the physical size of a pixel. Obviously, it depends on both the resolution of the recording procedure and the distance of the camera to the analyzed system. The farther the camera is located, the bigger the error.

To determine which part of the error is due to the sPGD procedure and which to experimental errors, the method was tested against a closed-form, error-free synthetic solution of the EBN model, taken as ground truth. We determined that the sPGD technique was responsible of around 1% of the global error.

For four different load positions, the maximum error at any of the white points in Figure 6 was, respectively, 3.00, 2.53, 2.39, and 2.76 mm. The average error was however, for these same cases, 0.658, 0.753, 0.805, and 0.775 mm, showing the great precision of the proposed technique.

7 | CONCLUSIONS

In this article we have proposed a new type of digital twin. As distinct features, the proposed technique incorporates computer vision as the experimental setup, and augmented reality toward the ease of interpretation of the information provided by the model, whose parameter values are determined by means of data assimilation procedures, combined with model reduction techniques. These are indispensable to obtain real-time performance at video rates (30–60 frames per second).

But, notably, the salient feature of the digital twin relies in its ability to correct itself in an unsupervised manner, by detecting systematic differences between measurements and predictions. If this is the case, the model is corrected by employing the same measurements obtained by the camera. To this end, a sparse-proper generalized decomposition approach has been chosen. It allows to construct the model correction in a separated form. This separated form is especially convenient for the real-time solution of the inverse problem.

The proposed methodology has demonstrated to obtain very accurate results, always bounded by experimental errors. This opens the possibility of defining a new generation of digital twins, which we have coined as hybrid twins.

Recent advances in the field of scientific machine learning open the possibility to extend these result and go much further. However, there remains a wide field to explore. For instance, how to detect more intricate features in the models, such as internal variables, for instance. This constitutes the goal of our current efforts of research.

ACKNOWLEDGMENTS

This project has been partially funded by the ESI Group through the ESI Chair at ENSAM Arts et Metiers Institute of Technology, and through the project “Simulated Reality” at the University of Zaragoza. The support of the Spanish Ministry of Economy and Competitiveness through grant number CICYT-DPI2017-85139-C2-1-R and by the Regional Government of Aragon and the European Social Fund, are also gratefully acknowledged.

ORCID

Eliás Cueto  <https://orcid.org/0000-0003-1017-4381>

REFERENCES

1. Shafto M, Conroy M, Doyle R, et al. Modeling, simulation, information technology & processing roadmap. *Nat Aeronaut Space Administr.* 2012.
2. El Saddik A. Digital twins: the convergence of multimedia technologies. *IEEE MultiMed.* 2018;25(2):87-92. <https://doi.org/10.1109/MMUL.2018.023121167>.
3. Various Authors. *Final Report. DDDAS Workshop.* Arlington, VA: National Science Foundation; 2006.
4. Blasch E, Seetharaman G, Reinhardt K. Dynamic data driven applications system concept for information fusion. *Proc Comput Sci.* 2013;18:1999-2007. doi:10.1016/j.procs.2013.05.369
5. Douglas C, Efendiev Y, Ewing R, Ginting V, Lazarov R. dynamic data driven simulations in stochastic environments. *Computing.* 2006;77(4):321-333. <https://doi.org/10.1007/s00607-006-0165-3>.
6. Darema F. Grid computing and beyond: the context of dynamic data driven applications systems. *Proc IEEE.* 2005;93(3):692-697. <https://doi.org/10.1109/JPROC.2004.842783>.
7. Darema F. *Dynamic Data Driven Applications Systems (DDDAS).* Arlington, VA: Air Force Office Of Scientific Research; 2013.
8. Lee EA. Cyber physical systems: design challenges. Paper presented at: Proceedings of the 2008 11th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing (ISORC); 2008:363-369; New York, NY.
9. Ghnatios C, Masson F, Huerta A, Leygue A, Cueto E, Chinesta F. Proper generalized decomposition based dynamic data-driven control of thermal processes. *Comput Methods Appl Mech Eng.* 2012;1:213-216. <https://doi.org/10.1016/j.cma.2011.11.018>.
10. Gonzalez D, Masson F, Poulhaon F, Cueto E, Chinesta F. Proper generalized decomposition based dynamic data driven inverse identification. *Math Comput Simul.* 2012;82:1677-1695.
11. Peherstorfer B, Willcox K. Dynamic data-driven reduced-order models. *Comput Methods Appl Mech Eng.* 2015;291:21-41. <https://doi.org/10.1016/j.cma.2015.03.018>.
12. Allaire D, Kordonowy D, Lecerf M, Mainini L, Willcox K. Multifidelity DDDAS methods with application to a self-aware aerospace vehicle. *Proc Comput Sci.* 2014;29:1182-1192. <https://doi.org/10.1016/j.procs.2014.05.106>.
13. Cortial J, Farhat C, Guibas LJ, Rajashankar M. Compressed sensing and time-parallel reduced-order modeling for structural health monitoring using a DDDAS. In: Shi Y, Albada GD, Dongarra J, Sloot PMA, eds. *Computational Science – ICCS 2007.* Berlin, Heidelberg/Germany: Springer; 2007:1171-1179.
14. Peherstorfer B, Willcox K. Detecting and adapting to parameter changes for reduced models of dynamic data-driven application systems. *Proc Comput Sci.* 2015;51:2553-2562.
15. Chinesta F, Cueto E, Abisset-Chavanne E, Duval JL, El Khaldi F. Virtual, digital and hybrid twins: a new paradigm in data-based engineering and engineered data. *Arch Comput Methods Eng.* 2020;27:105-134. <https://doi.org/10.1007/s11831-018-9301-4>.
16. Hey T, Tansley S, Tolle K. *The Fourth Paradigm: Data-Intensive Scientific Discovery.* Microsoft Research; 2009.
17. Lee S, Baker N. *Basic Research Needs for Scientific Machine Learning: Core Technologies for Artificial Intelligence.* USDOE Office of Science (SC); 2018.
18. Lu L, Meng X, Mao Z, Karniadakis GE. DeepXDE: a deep learning library for solving differential equations; 2019. arXiv e-prints. ;arXiv:1907.04502.
19. Biocca Frank. Chapter 3 Intelligence augmentation: the vision inside virtual reality. In: Gorayska Barbara, Mey Jacob L., eds. *Cognitive Technology, Advances in Psychology*, vol. 113: North-Holland 1996 (pp. 59 - 75).
20. Daily M, Oulasvirta A, Rekimoto J. Technology for Human Augmentation. *Computer.* 2017;50(2):12-15.
21. Engelbart DC, & Friedewald M. *Augmenting Human Intellect: A Conceptual Framework.* Stanford Research Institute; 1997.
22. Michopoulos JG, Hermanson JC, Furukawa T. Towards the robotic characterization of the constitutive response of composite materials. *Compos Struct.* 2008;86(1-3):154-164.
23. Michopoulos JG, Furukawa T. Towards hierarchical design optimization for simultaneous composite material characterization and adjustment of the corresponding physical experiments. *Inverse Probl Sci Eng Formerly Inverse Probl Eng.* 2008;16(6):763-775.
24. Badiás A, Alfaro Iciar, González D, Chinesta F, Cueto Eliás. Reduced order modeling for physically-based augmented reality. *Comput Methods Appl Mech Eng.* 2018;341:53-70. <https://doi.org/10.1016/j.cma.2018.06.011>.
25. Chinesta F, Cueto E. *PGD-Based Modeling of Materials, Structures and Processes.* Switzerland: Springer International Publishing; 2014.
26. Chinesta F, Keunings R, Leygue A. *The Proper Generalized Decomposition for Advanced Numerical Simulations.* Switzerland: Springer International Publishing; 2014.

27. Cueto E, González D, Alfaro I. *Proper Generalized Decompositions: An Introduction to Computer Implementation with Matlab*. Springer Briefs in Applied Sciences and Technology. New York, NY: Springer International Publishing; 2016.
28. Rozza G, Huynh DBP, & Patera AT. Reduced Basis Approximation and a Posteriori Error Estimation for Affinely Parametrized Elliptic Coercive Partial Differential Equations. *Arch Computat Methods Eng*. 2008;15:229. <https://doi.org/10.1007/s11831-008-9019-9>
29. Hesthaven J, Rozza G, Stamm B. *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*. New York, NY: Springer Verlag; 2015.
30. Rozza G. Fundamentals of reduced basis method for problems governed by parametrized PDEs and applications. In: Ladeveze P, Chinesta F, eds. *CISM Lectures Notes "Separated Representation and PGD Based Model Reduction: Fundamentals and Applications"*. New York, NY: Springer Verlag; 2014.
31. González D, Badiás A, Alfaro Iciar, Chinesta F, C Elías. Model order reduction for real-time data assimilation through Extended Kalman Filters. *Comput Methods Appl Mech Eng*. 2017;326(Suppl C):679-693. <https://doi.org/10.1016/j.cma.2017.08.041>.
32. Mori H. Transport, collective motion, and Brownian motion. *Prog Theor Phys*. 1965;33(3):423-455. <https://doi.org/10.1143/PTP.33.423>.
33. Zwanzig R. Nonlinear generalized Langevin equations. *J Stat Phys*. 1973;9(3):215-220.
34. Chinesta F, Cueto E, Grmela M, Moya B, Pavelka M. Learning physics from data: a thermodynamic interpretation; 2019. arXiv preprint arXiv:1909.01074.
35. Ibáñez R, Abisset-Chavanne E, Ammar A, et al. A multidimensional data-driven sparse identification technique: the sparse proper generalized decomposition. *Complexity*. 2018;2018.
36. Badiás A, Alfaro I, Gonzalez D, Chinesta F, Cueto E. Real-time interaction of virtual and physical objects in mixed reality applications. *Comput Mech*. 2020;121:3849–3868.
37. Badiás A, Curtit S, González D, Alfaro Iciar, Chinesta F, Cueto Elías. An augmented reality platform for interactive aerodynamic design and analysis. *Int J Numer Methods Eng*. 2019;120(1):125-138. <https://doi.org/10.1002/nme.6127>.
38. Bradski G, Kaehler A. *Learning OpenCV: Computer vision with the OpenCV library*. O'Reilly Media, Inc; 2008.
39. Shi J, Tomasi C. *Good Features to Track*. Seattle, WA, USA: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition; 1994:593-600.
40. Badiás A, González D, Alfaro I, Chinesta F, Cueto E. Local proper generalized decomposition. *Int J Numer Methods Eng*. 2017;112(12):1715, nme.5578-1732. <https://doi.org/10.1002/nme.5578>.
41. Ibáñez R, Abisset-Chavanne E, Chinesta F, Huerta A, Cueto Elías. A local multiple proper generalized decomposition based on the partition of unity. *Int J Numer Methods Eng*. 2019;120(2):139-152. <https://doi.org/10.1002/nme.6128>.
42. Chinesta F, Ammar A, Cueto E. Recent advances in the use of the Proper Generalized Decomposition for solving multidimensional models. *Arch Comput Methods Eng*. 2010;17(4):327-350.
43. Manzoni A, Quarteroni A, Rozza G. Computational reduction for parametrized PDEs: strategies and applications. *Milan J Math*. 2012;80:283-309.
44. Ammar A, Huerta A, Chinesta F, Cueto Elías, Leygue A. Parametric solutions involving geometry: a step towards efficient shape optimization. *Comput Methods Appl Mech Eng*. 2014;268:178-193. <https://doi.org/10.1016/j.cma.2013.09.003>.
45. Chinesta F, Leygue A, Bordeu F, et al. PGD-based computational vademecum for efficient design, optimization and control. *Arch Comput Methods Eng*. 2013;20(1):31-59. <https://doi.org/10.1007/s11831-013-9080-x>.

How to cite this article: Moya B, Badiás A, Alfaro I, Chinesta F, Cueto E. Digital twins that learn and correct themselves. *Int J Numer Methods Eng*. 2022;123:3034–3044. <https://doi.org/10.1002/nme.6535>