

ARTICLE TYPE

An Augmented Reality platform for interactive aerodynamic design and analysis[†]

Alberto Badías¹ | Sarah Curtit^{1,2} | David González¹ | Icíar Alfaro¹ | Francisco Chinesta³ | Elías Cueto¹

¹Aragon Institute of Engineering Research, Universidad de Zaragoza, Edificio Betancourt. Maria de Luna, s.n., 50018 Zaragoza, Spain

²Student, École nationale des ponts et chaussées, 6-8 Avenue Blaise Pascal, 77420 Champs-sur-Marne, France

³ESI Group Chair @ PIMM - Procédés et Ingénierie en Mécanique et Matériaux, ENSAM ParisTech, 151 boulevard de l'Hôpital, Paris, France

Correspondence

Elías Cueto. Email: ecueto@unizar.es

Summary

While modern CFD tools are able to provide the user with reliable and accurate simulations, there is a strong need for interactive design and analysis tools. State of the art CFD software employs massive resources in terms of CPU time, user interaction and also GPU time for rendering and analysis. In this work we develop an innovative tool able to provide a seamless bridge between artistic design and engineering analysis. This platform has three main ingredients: computer vision to avoid long user interaction at the pre-processing stage, machine learning to avoid costly CFD simulations and augmented reality for an agile and interactive post-processing of the results.

KEYWORDS:

Augmented reality, machine learning, real time, CFD, interactivity.

1 | INTRODUCTION

Very much like other aspects of our everyday life, simulation is living profound changes. It is not infrequent to hear about “democratisation”—more than deployment—of simulation or even about *appification*. In general, by *democratisation* we mean to make simulation accesible to users that previously did not feel the need for it, but now consider it interesting. However, many undoubtedly interesting applications of simulation in engineering practice come at the price of many hours of user interaction followed by many CPU hours to obtain results. And this results in a barrier for many potential users.

In industry, particularly in automotive manufacturing, there is a claim for interactive ways to connect designers (artists, in sum) and engineers [1]. For this to be possible, developing real-time, interactive simulation apps that could provide immediate responses on the physical effects of purely artistic decisions on mockups, for instance, could be of great help. This immediacy and interactivity should include the pre-processing stage in which engineers construct the CAE model of the product, the meshing operation for the development of a FEA model and a reasonable time to response.

In parallel, with the global deployment of mobile devices (smartphones, tablets), there is a tendency towards simulation *appification*, i.e., the implementation on such platforms of the usual simulation tools. While the performance of these platforms is nowadays impressive, they are undoubtedly less powerful than usual workstations. This imposes additional requirements to any realistic attempt to simulation democratization.

Augmented reality (AR) has gained popularity given the realism and ease of interpretation of the information it provides. Therefore, AR can now be seen as a feasible way for simulation post-processing. However, AR makes intensive use of computer vision techniques [2] [3] and therefore induces additional restrictions to the simulation pipeline [4].

[†]This project has been partially funded by the ESI Group through the ESI Chair at ENSAM ParisTech and through the project “Simulated Reality” at the University of Zaragoza. The support of the Spanish Ministry of Economy and Competitiveness through grant number DPI2015-72365-EXP and by the Regional Government of Aragon and the European Social Fund, are also gratefully acknowledged.

In this work we present our first attempt to democratize simulation by developing a platform that copes with the aforementioned requirements. We try to directly suppress the user time employed in the construction of the model. Instead, it is required that the model comes directly from physical reality, through computer vision techniques. As a proof of concept, we focus on car aerodynamics, and thus the geometry of the car (or prototype) will be captured by a commodity camera. No stereo or RGBD cameras are needed, in principle, even if their use is obviously possible and could eventually ease and improve the accuracy of the process. Previous approaches to this problem employ this type of cameras (Microsoft Kinect cameras, for instance, in [5], that capture depth measures by means of an infrared laser).

Geometry of the objects to analyze will thus be acquired by computer vision, thus bypassing the need for time-consuming CAE pre-processing steps. Usually, the same device that captures geometry will serve to depict the results in an AR framework, but many different possibilities exist. Our approach makes use of feature-based Simultaneous Localization and Mapping (SLAM) techniques to acquire the geometry of the object with a single monocular camera, even though many modern smartphones are already equipped with stereoscopic cameras. In particular, we employ ORB-SLAM techniques [6]. These techniques allow us to capture selected points on the surface of the object, whose geometry will need nevertheless to be reconstructed from these points. In other words, a suitable parametrization of the shape of the object needs to be constructed. Previous works include the use of free-form deformation techniques [7] or the closely-related approach in [8].

Our approach to the problem is data-driven [9] [10] [11] [12] [13] [14]. This means that we employ manifold learning techniques [15] over a data base of previously computed CFD results. While there is a vast corps of literature on (linear) model reduction of flow problems [16] [17] [18] [19] [20] [21] [22], these techniques rarely achieve real-time performance. For this particular application, by real-time we mean a code able to provide feedback response at some 30–60 Hz (i.e., the usual number of frames per second in modern smartphone cameras).

Other approaches, however, achieve interactivity by means of machine learning [8] or by model reduction [1], but do not consider augmented reality output. On the other side of the spectrum, some previous studies have considered the *brute force* approach [5]. This type of approach makes intensive use of GPUs capabilities in a Lattice Boltzmann framework and are not well suited, therefore, for their implementation on a mobile platform.

The employ of machine learning towards learning physical dynamics has gained considerable interest in recent times [28] [29] [30]. In particular, deep learning architectures are of course tempting [23]. In general, while model order reduction methods [24][25] achieve important speedup by means of a considerable reduction in the number of degrees of freedom and also provide with important insight on the physics of the problem, deep learning architectures perform in nearly the opposite way [26]. They are able to capture multiscale phenomena or transient features and, at the same time, are able to capture invariants of the system, but need for a tremendous effort to train the neural networks and, conversely, provide little or no interpretation of the results. Some scientists have warned about the possible implications that the employ of machine learning on top of biased data could have [27].

Closely related, there is a strong research activity in the development of physics-aware artificial intelligence—in other words, machine learning approaches that satisfy first principles by construction [31] [32] [33]. Our approach here is to employ, so to speak, the best of both worlds. Non-linear dimensionality reduction techniques—in particular, manifold learning—are able to substantially reduce the number of degrees of freedom of the problem, very much like classical, linear model order reduction. But, at the same time, they provide very useful insight on the non-linearity of the problem. As will be noticed from the results here presented, manifold learning techniques provide clear physical explanations on the structure of data, reduce the number of degrees of freedom and their dimensionality, and allow for very efficient computational approaches.

With all these ingredients—computer vision, machine (manifold) learning and augmented reality—we construct a reliable simulation platform that, even if we admit errors higher than those typical of state-of-the-art high-fidelity simulations, provide the user with a unique degree of interactivity and immediacy.

The paper is structured as follows. In Section 2 we provide an overall description of the method we have developed. In Section 3 we give detailed insight on the construction of the database and its subsequent embedding in a lower dimensional space. Then, in Section 4 we briefly describe the state-of-the-art computer vision techniques employed to capture cars not being in the database. In Section 5 we describe how to interactively depict CFD results on top of the video stream of the particular car of interest. Finally, in Section 6 we analyze the accuracy of the results, computational cost, etc. The paper is completed with some final comments about the interest of the proposed technique.

2 | OVERALL DESCRIPTION OF THE METHOD

Figure 1 depicts a sketch of the general pipeline of the method. The key ingredient is the employ of manifold learning techniques to *learn* car aerodynamics. This will allow us to avoid simulating in real time. Instead, we *interpolate*—in the right manifold space—existing CFD results (represented in the left column of Fig. 1). By obtaining a low-dimensional embedding of these results we will not only ensure to interpolate in the right space (instead of the Euclidean one) but also take profit of the non-linear reduced order modeling capabilities of Locally Linear Embedding techniques [34], which will be thoroughly described in Section 3.

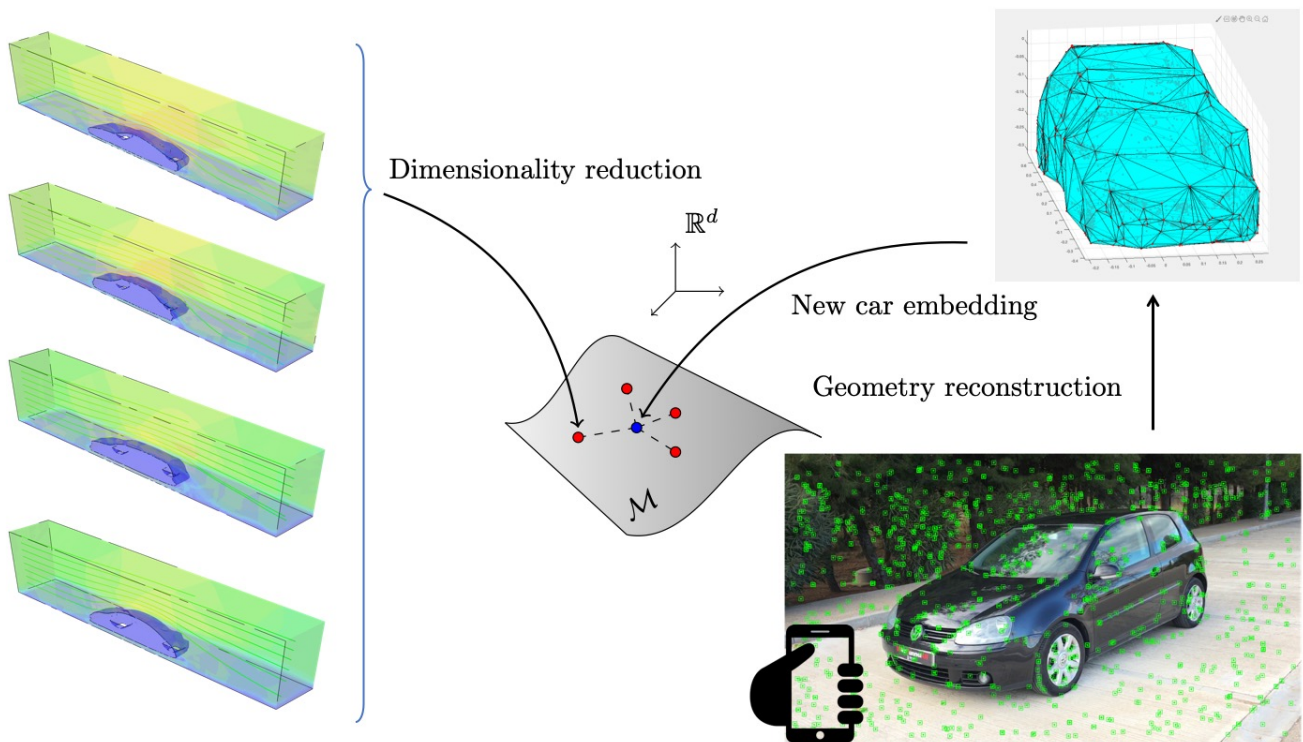


FIGURE 1 Schematic description of the method. We first construct a database of CFD results for known car geometries (left column). These results are embedded onto a low-dimensional manifold \mathcal{M} by means of LLE techniques. Once a new car—not in the database—is captured with the camera, the features are first processed so as to obtain an approximate envelope of the new car geometry. This new geometry is then embedded onto the manifold (blue point) and its CFD results interpolated among its neighbors in the manifold (red points).

These results are obtained by usual CFD methodologies, starting from CAD descriptions of car bodies. Once unveiled, the manifold structure of these data, \mathcal{M} , will allow us to easily incorporate new designs, not previously analyzed (right column of Fig. 1). With the help of a standard smartphone we will capture the bodywork of a new car or prototype. To this end, it is necessary to film around the car, so as to capture its entire geometry. SLAM techniques [6] will allow us to detect special points in the body of the car called ORB features (green points in the bottom-right corner of Fig. 1) [35].

After a reconstruction of the body of the car from these ORB points (top right corner of Fig. 1) we are in the position of embedding this new design onto the already available manifold of shapes \mathcal{M} . This will also provide us with the necessary weights so as to interpolate from surrounding CFD results in the database. This simple approach will allow us to obtain velocity maps, streamlines, etc. without the burden of meshing procedures, long CPU computations and subsequent post-processing.

Once the new results have been obtained, they are ready to be represented on the screen, on top of the video stream. For this to be possible, a *registration* process will be needed. By image registration we mean the process of finding the common system

of reference between the video frames and the just obtained CFD results. So to speak, we need to *align* both so as to construct an augmented reality post-processing of these results.

These steps are detailed in the following sections. We begin by describing the steps towards the obtention of the car database.

3 | CONSTRUCTION OF A SUITABLE DATABASE OF CFD RESULTS

3.1 | Parameterization of the shape

A crucial step in the machine learning of car aerodynamics is that of parameterizing shape. This issue has been deeply analyzed in [8], for instance. In the reduced order modeling literature there is another vast corps of literature, see, for instance, [20] [7] [36] [17] [37] for works employing freeform deformation techniques. Of course, usual CAD techniques based on the employ of NURBS or related approaches cannot be employed here, since the capture of shapes by computer vision is obliged in the tool we are developing. This problem shares many similarities, for instance, with the construction of patient avatars in the medical community [38] [39] [40].

The original database was composed by object (*.obj) files. Prior to converting them to a manageable format, compatible with the computer vision system, a de-refinement of the bodywork of every analyzed model is mandatory, see Fig. 2 . We employed the same de-refinement method as in [8]. In the examples analyzed herein, the intermediate level of refinement was chosen, that corresponds with the center figure in Fig. 2 . Of course, the final accuracy of the method strongly depends on the level of detail to which the body of the car is represented. Further refinement leads of course to higher levels of accuracy.

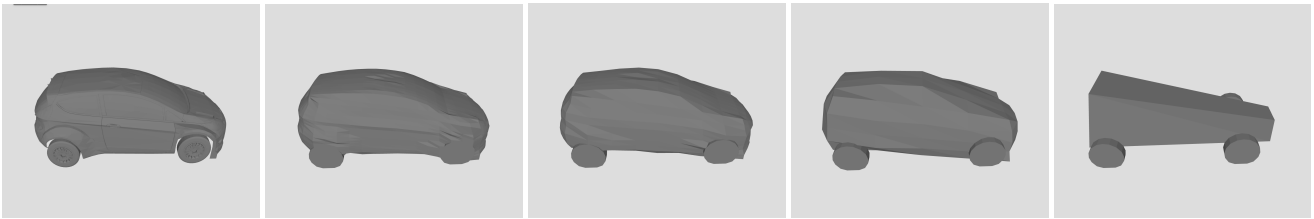


FIGURE 2 Model de-refinement for the Ford Fiesta example.

This de-refined geometry is then embedded into a grid of $140 \times 40 \times 20$ points uniformly distributed on a volume of $14 \times 5 \times 4$ cubic meters around each half car (we exploit here the symmetry of cars so as to include only one half of the bodywork in the database). While in some of the previous works of the authors [38] we employed a signed distance field to characterize the shape, in this work we have found that the employ of a presence function

$$\phi(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in \Omega, \\ 0 & \text{if } \mathbf{x} \notin \Omega, \end{cases} \quad (1)$$

where Ω represents the boundary of the car, rendered better results for the manifold learning procedure to be described next. In other words, we characterized each particular shape by a vector $\mathbf{z} \in \mathbb{R}^D$, where $D = 140 \times 40 \times 20 = 112000$, whose entries are the values of $\phi(\mathbf{x}_j)$ at each node of the grid, \mathbf{x}_j , $j = 1, \dots, 112000$.

3.2 | Manifold learning

The key ingredient in our method is to hypothesize that car shapes describe a manifold structure whose precise form is to be found. To this end, we employ Locally Linear Embedding (LLE) techniques [34]. In essence, LLE assumes that, given the fact that a manifold is homeomorphic to a flat space in a sufficiently small neighborhood of each point, each datum can be reasonably well approximated through a linear combination of its nearest neighbors—whose precise number M is to be fixed by the user—. Therefore, LLE assumes that for each \mathbf{z}_m a linear reconstruction exists of the form

$$\mathbf{z}_m = \sum_{i \in S_m} W_{mi} \mathbf{z}_i, \quad (2)$$

where W_{mi} are the unknown weights and S_m the set of the M -nearest neighbors of z_m .

The search for the weights is done by minimizing the functional

$$\mathcal{F}(W) = \sum_{m=1}^M \left\| z_m - \sum_{i=1}^M W_{mi} z_i \right\|^2,$$

where W_{mi} is zero if $z_i \notin S_m$, i.e., if cars i and m are not neighbors in the manifold of shapes.

LLE also hypothesizes that an embedding exists onto a lower-dimensional space \mathbb{R}^d , with $d \ll D$, such that the manifold structure is conserved. In other words, a new set of coordinates $\xi \in \mathbb{R}^d$ exists for each car $z \in \mathbb{R}^D$ and they can be found by minimizing a new functional

$$\mathcal{G}(\xi_1, \dots, \xi_M) = \sum_{m=1}^M \left\| \xi_m - \sum_{i=1}^M W_{mi} \xi_i \right\|^2,$$

where the neighboring relation between points (i.e., the weights W_{mi}) are maintained. See Fig. 3 for a description of LLE.

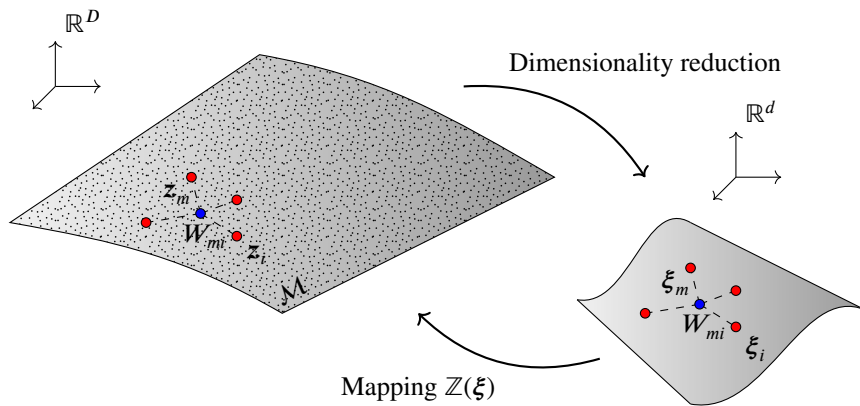


FIGURE 3 Hypothesis about the existence of a manifold \mathcal{M} on which the car shapes live. The small black dots represent the car database in a high-dimensional space \mathbb{R}^D . These data are then embedded onto a lower-dimensional manifold \mathbb{R}^d , with $d \ll D$. A new car not in the database (blue dot) will thus be interpolated from its neighboring cars (red points) through the weights W_{mi} provided by LLE, that define a mapping $Z(\xi)$. By hypothesis, LLE assumes that the same neighborhood relationships hold in the high-dimensional manifold \mathcal{M} .

3.3 | Our database

We analyzed a set of 80 different car bodies. All of them are subjected to a CFD analysis using commercial software. Fig. 4 shows a typical mesh of the environment of one of the models. All of them are subjected to a uniform velocity of 10 m/s at the leftmost part of the meshed volume. This gives us a Reynolds number on the order of 9000.

After the CFD computations, a projection of the velocity field onto the same grid mentioned in the Section 3.1 is made for subsequent interpolation purposes. Our cars are then manually classified into four different groups: large cars (Audi Q7, Hummer, Jeep Grand Cherokee, VW Touareg, among others), sport cars (Mercedes AMG GTS, BMW Z4, Audi R8, BMW i8, Porsche 911, ...), old cars (Pontiac firebird, Dodge Charger, Chevrolet Camaro, ...) and a fourth set of “various”, which comprises not easily classifiable models.

Our experiments showed that a range between $M = 6 - 15$ neighbors provided a good embedding onto \mathbb{R}^2 . Figure 5 shows the embedding obtained with $M = 9$ neighbors, where it can be noticed how LLE clearly clusterizes sport cars apart from large, familiar cars, for instance. Old cars, whose shape does not follow any specific pattern, and those cars in the “various” group are somewhat mixed in the embedding procedure, as expected.

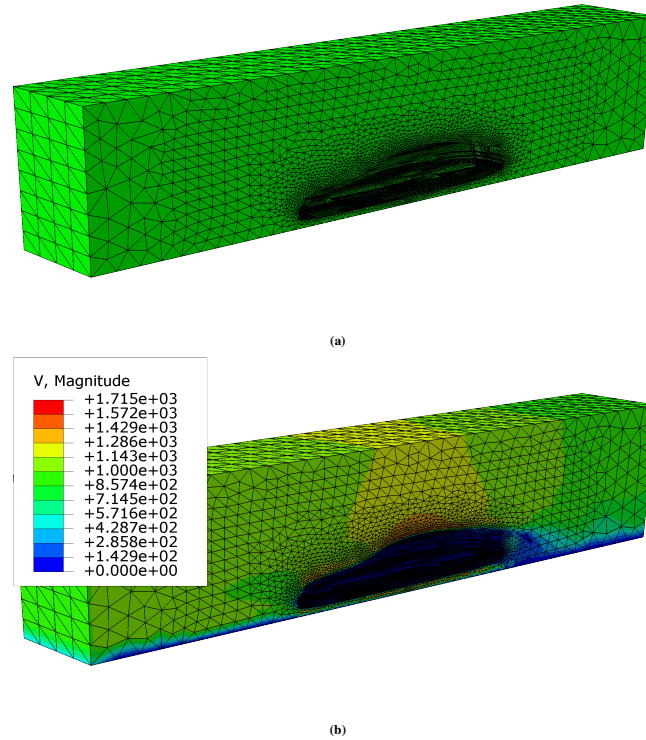


FIGURE 4 (a) Typical FE mesh employed in the CFD calculations. (b) Contour plot of the velocity module for the same model.

4 | COMPUTER VISION FOR NEW CAR MODELS

While the previous sections described the off-line work to construct the car database, in the following section we describe the on-line work done for the characterization of the new geometry and the subsequent interpolation of the results on the manifold of shapes.

After the database has been completed, and its dimensionality conveniently reduced, the pipeline of our platform faces the problem of reconstructing the geometry of a new car. In general, a new car or a new design will not be in the database, so we assume that no CAD description is available. Instead, the user will make use of a commodity camera (through a smartphone, for instance) to capture the new geometry.

4.1 | Capturing the geometry from a video sequence

The problem of reconstructing a three dimensional geometry from a set of two-dimensional images is known in the robotics field as the *Structure from Motion*, SfM, problem [2] [41] [42]. In essence, if we assume that the object to reconstruct is rigid—for the deformable case refer to [4] and references therein—the problem is sketched in Fig. 6 . The object is assumed to be attached to a fixed position in the frame of reference. The problem thus begins by determining the *pose*—location and orientation—of the camera.

The input image space is denoted by $\mathcal{I}_k \subset \mathbb{R}^2$, with $k = 1, \dots, n_{\text{images}}$, the number of available frames in the just captured video. We assume no batch approach in this platform but we do not consider the possibility of a video stream. In other words, the video has a finite number of frames.

Cameras perform a projection from three-dimensional, physical space to a two-dimensional pixel space. This projection is characterized by an *a priori* unknown camera operator denoted by $\Pi : \partial^* \Omega_t \rightarrow \mathcal{I}$ whose determination is known as the *calibration* of the camera. This projection can be expressed as $\Pi = \mathbf{K}[\mathbf{R}|t]$, where \mathbf{K} corresponds to the camera intrinsic matrix

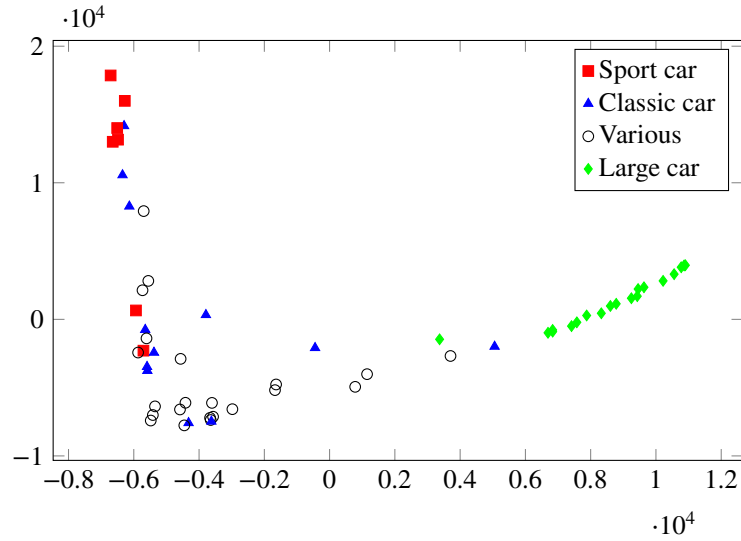


FIGURE 5 Results of the embedding of car geometries onto \mathbb{R}^2 by employing nine neighbors for each car. It can be noticed how LLE clusterizes sport cars far apart familiar (or “large”) cars, while old cars and other models without a distinctive characteristic are somehow mixed.

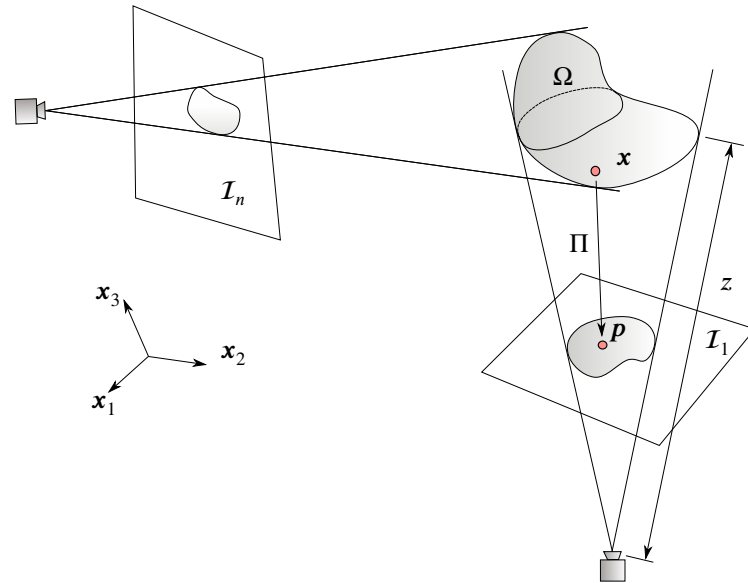


FIGURE 6 Sketch of the Structure from Motion problem. While the object is assumed to be rigid and attached to a fixed position in the frame of reference, the camera needs to move so as to reconstruct its three-dimensional geometry. Here, two video frames, namely \mathcal{I}_1 and \mathcal{I}_n are shown, together with the camera projections of the solid. The projection Π of a particular point \mathbf{x} on the visible surface of the solid gives rise to point \mathbf{p} in the pixel space.

and $[\mathbf{R}|\mathbf{t}]$ represents the camera extrinsic matrix. We denote $\mathbf{x} = (x, y, z)^\top$ and $\mathbf{p} = (p_1, p_2)^\top$ and a scale factor s , so that,

$$\begin{bmatrix} p_1 \\ p_2 \\ s \end{bmatrix} = \lambda \begin{bmatrix} f/d_x & 0 & c_x \\ 0 & f/d_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \underbrace{[\mathbf{R} \ \mathbf{t}]}_{\mathbf{x}^c} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}, \quad (3)$$

where \mathbf{x}^c represents the coordinates of \mathbf{x} in the camera system (a moving frame of reference attached to the camera, not the world reference frame). $\partial^*\Omega$ denotes the visible part of the boundary of the object of interest, Ω .

Two transformations exist: first, the camera extrinsic matrix transforms three-dimensional physical points in the world system of reference into three-dimensional points in the camera moving frame of reference. This transformation is equivalent to the composition of a rotation (\mathbf{R}) and translation (\mathbf{t}) between the corresponding points. On the other hand, the camera intrinsic matrix \mathbf{K} contains information about the projection centre (c_x, c_y) , the pixel size (d_x, d_y) and the focal distance, f , that maps three-dimensional camera points into two-dimensional pixel coordinates. Finally, λ represents a scale factor that takes into account the fact that, with a single picture, an indeterminacy exists that makes it impossible to discern between the projection of a big object positioned far from the objective or a small one at a closer position.

Therefore, to determine the camera pose and the location of a rigid object we need more than one single image. To solve this indeterminacy, SfM algorithms minimize the *reprojection error*. In other words, SfM minimize the geometric error corresponding to the image distance between a projected point and a measured one [3]. As sketched in Fig. 6, the camera performs a projection of the form

$$\mathbf{p} = \Pi(\mathbf{x}),$$

with $\mathbf{p} \in \mathbb{R}^2$ and $\mathbf{x} \in \mathbb{R}^3$. Note that we do not know the exact value of \mathbf{x} . Well established geometrical analysis will provide an estimation $\hat{\mathbf{x}}$ of the true position of the point [43] [3]. If we re-project this estimate for each frame by

$$\hat{\mathbf{p}} = \Pi(\hat{\mathbf{x}}),$$

the sought reprojection error takes the form $d(\mathbf{p}, \hat{\mathbf{p}})$.

This analysis is often made with respect to some marked points usually known as *fiducial markers*. In general, this analysis is greatly simplified if the object to reconstruct can be labelled at some points with the help of some visible sticker, for instance. However, for our purpose this is not feasible. The tracked points \mathbf{x} must be selected by the algorithm itself. To this end, we employ a method based on *features*, the so-called Oriented FAST and rotated BRIEF (ORB) method [35]. In fact, we employ the so-called ORB-based Simultaneous Localization and Mapping (ORB-SLAM) method, able to solve the SfM problem while at the same time it constructs a map of the environment [6]. These ORB features are the green points shown in Fig. 7.

4.2 | Projection onto the database

The ORB features just captured constitute a cloud of points in the boundary of the sought geometry. To determine the best possible approximation to the sought geometry we employ α -shapes [44]. Essentially, α -shapes reconstruct a volume given a cloud of points on its boundary with the help of one user-defined parameter, α , that represents the coarsest level of detail in the reconstructed geometry. Basically, α is on the order of the distance between recognized ORB features.

Of course, a stereo or RGBD camera would provide a much more detailed geometry representation. We analyze, however, the case of deployed hardware, and consider that a simple smartphone or tablet is enough to obtain a reasonable representation of each car bodywork. With the just reconstructed geometry, we embed it in the same grid employed to construct the car database and evaluate the presence function $\phi(\mathbf{x})$ at each of the nodes of the grid, \mathbf{x}_j , $j = 1, \dots, D$, see Fig. (9). This will provide a new vector $\mathbf{z}^{\text{new}} = [\phi(\mathbf{x}_1) \dots \phi(\mathbf{x}_D)]^\top$ for the geometry at hand.

This new vector \mathbf{z}^{new} can easily be embedded on the manifold without the need of re-computing every weight, by invoking incremental LLE algorithms, see [45], for instance. This will provide the new weights W_i , $i = 1, \dots, M$.

The key ingredient in our method is the assumption that the velocity field (for sufficiently low Re numbers) of a new car can be obtained in an accurate enough manner by linear interpolation among the neighboring models in the manifold. In other words, that the new velocity field is approximated by

$$\mathbf{v}^{\text{new}}(\mathbf{x}_j) \approx \sum_{i=1}^M W_i \mathbf{v}_i(\mathbf{x}_j), \quad (4)$$

with $\mathbf{v}_i(\mathbf{x}_j)$ the velocity field for the i -th neighboring car model at the j -th node of the grid. This assumption, which may seem too restrictive, has demonstrated already to provide reasonably accurate results in a completely different context, but in a similar procedure, see [38].

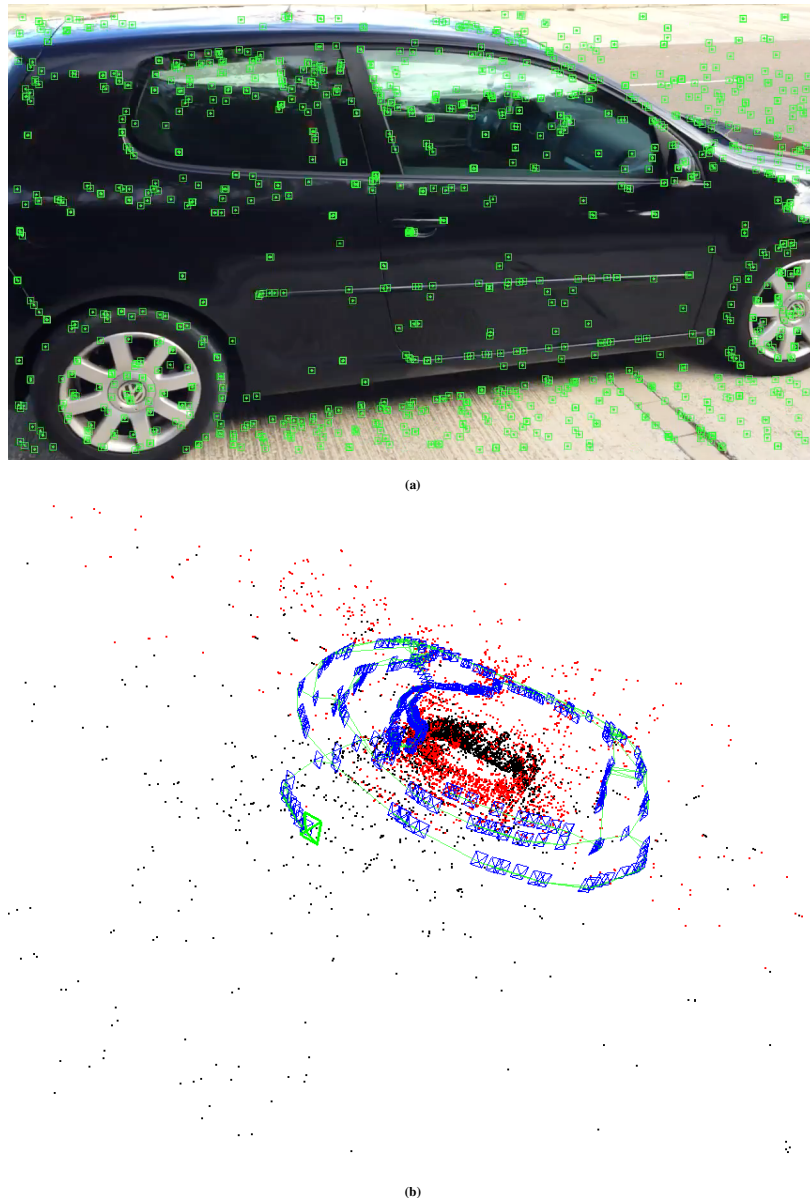


FIGURE 7 (a) An example of the captured ORB features for the VW Golf model. (b) A reconstruction of the camera movements in the video. Blue pyramids indicate the position and orientation (pose) of the camera. Dots represent the ORB's detected by the SLAM methodology, that belong both to the car and its surroundings.

5 | AUGMENTED REALITY OUTPUT

5.1 | Car identification

ORB-SLAM is probably the best technique to reconstruct static 3D scenes in a sparse way, if we think of employing a commodity camera. We can estimate the 3D position $\mathbf{x}(x, y, z)$ of the most relevant points of a scene, extracted by using the feature descriptor ORB. This means that we need a relatively low computational power to work, it does not require the use of graphical acceleration by GPU, at the cost of reducing the number of observed points. With this system we do not expect to obtain a dense cloud of points, as it happens with other acquisition systems (such as RGB-D devices or stereo systems), but to be able to take very precise measurements in real time with little computational cost. It is also important to point out that densification techniques can be applied a posteriori to obtain a denser 3D point cloud, such as [46].

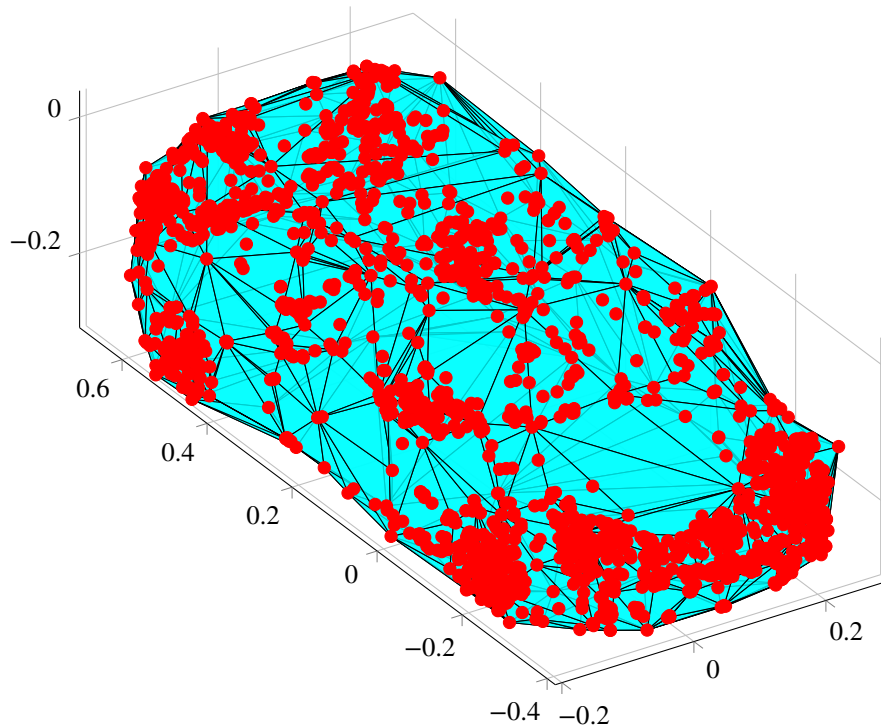


FIGURE 8 Reconstruction of the VW Golf model. Red points represent to detected ORB features. The blue volume is reconstructed by an α -shape algorithm.

Since we are using a monocular system (like the camera we can have in our smartphone) we can not recover the scale of the model without knowing any of the real dimensions of the objects (an RGB-D system or stereo camera could give us this information, but a standard user does not usually have this type of technology). Nor can we identify the direction of gravity (it could be done with an *Inertial Measurement Unit* or IMU), which complicates the estimation of the location or *pose* of the car in space. Our objective is to depend only on the visual system, obtaining the information of the scene from only a monocular camera.

We start, therefore, from a set of three-dimensional points estimated in the scene, where we need to identify the geometry of a car. For this we assume that the user has surrounded a real car with its monocular system, where the car appears relatively centered in the cloud of points. Figure 10 collects the algorithm used to identify, segment and locate the position of the car from the point cloud.

The first task of the algorithm is a point culling process depending on the local density surrounding any point of the cloud. We use this type of cleaning methods because sparse SLAM methods can add noise in the triangulation step. The use of a pyramid of image scales may also encourage this type of defects.

Next we estimate the centroid of the cleared point cloud, which allows us to locate where the most relevant object in the scene is located. After this, we eliminate those points that are further away from 95% of the distances of all points from the centroid. This allows us to work without the need to add adjustment parameters.

We then estimate the plane of the ground, because we assume that the car is located on a large plane with texture. So we estimate the main floor plane, and remove the points contained in that plane.

A principal direction analysis is then applied to estimate the predominant direction of the vehicle, that with the centroid and the gravity vector (estimated from the floor plane) allow us to fix the coordinates origin of the car.

The last step is the scaling. A noise removing process is again applied based in the histogram of positions to remove areas of isolated points. Then the car is inserted within a box to facilitate the estimation of scale and geometric transformations, which are then applied to move the car to the origin of coordinates and scale it with a unitary scale. We can do this because although we do not recover the scale, if the camera is well calibrated, the system is able to obtain a very precise *3D aspect ratio* between the

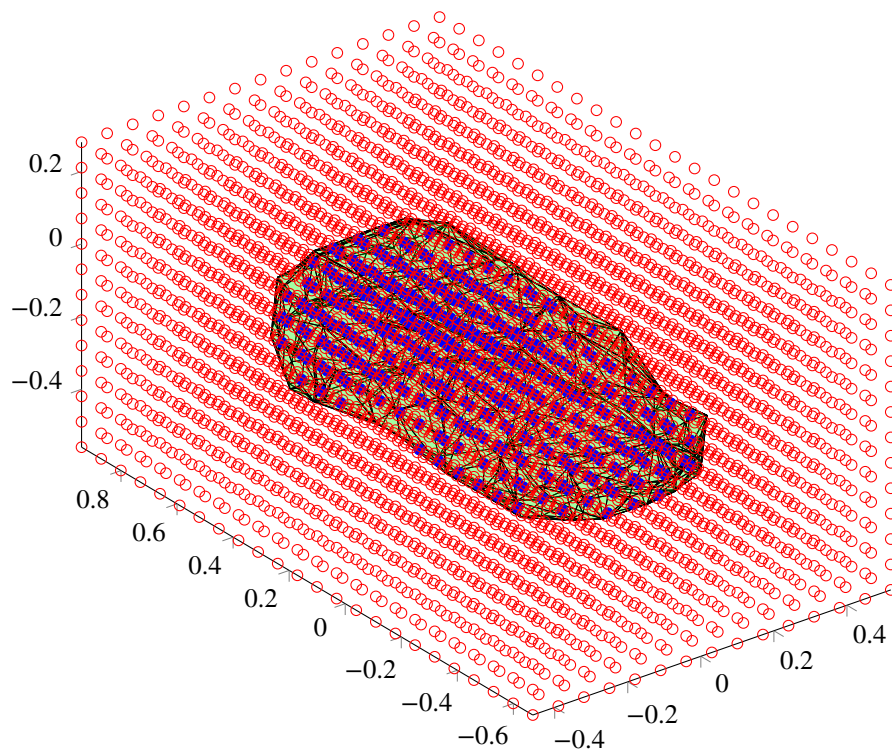


FIGURE 9 The car's geometry is then embedded on a grid and the presence function, Eq. (1), evaluated at each of its nodes. Red nodes are evaluated to zero (they fall outside of the car) while blue ones are evaluated to one (they fall inside). Note that the grid's resolution is lower than that actually employed ($140 \times 40 \times 20$) just to make the figure visible.

3 dimensions. In other words, even if the object is in another scale, moved and oriented in another direction, it is not necessary to apply different scales in the X, Y or Z axes, but a common scale of type $\mathbf{x}_{SC}(x, y, z) = \lambda(\mathbf{x}_{ORB}(x, y, z))$, where \mathbf{x}_{SC} are the 3D positions of the geometry with *scale correction* and $\mathbf{x}_{ORB}(x, y, z)$ are the initial 3D coordinates estimated by the SLAM system. This allows us to compare the estimated car with the database of the rest of the cars, where we have already precomputed the aerodynamic solution.

It is important to note that we could also have used modern techniques of recognition of car pose through images only, employing learning techniques [47], [48]. However, we think that it is possible that they do not reach the level of precision we expect for the geometry of a car. In addition, with our method we avoid the labeling and training phase. Working with image masks and then estimating the position of the 3D points by triangulation may involve similar noise problems, and some points could be located outside the real geometry of the car. Therefore, we should apply some filtering techniques again, as we have done in our proposed method. The tests we have done show that, although we use classic techniques in our method, they work in a robust way.

5.2 | Final appearance

The previous identification algorithm provides very valuable information for the final augmentation of the video. In particular, having the reconstructed three-dimensional geometry is a necessary step for the computation of occlusion culling in the results, i.e., disabling objects not seen by the camera to appear in the image. All the steps in the augmentation process are implemented in OpenGL [49]. The final appearance of the augmented results is shown in Fig. 11 .

```

function CarDetection (input SLAMCloud, output cleanCloud)
{
    densityPointCulling(SLAMCloud);
    centroid = geometryAnalysis(SLAMCloud);
    centroidPointCulling(SLAMCloud);
    floor = planeEstimationRANSAC(SLAMCloud);
    floorRemoving(SLAMCloud);
    PrincipalDirections = PrincipalComponentAnalysis(SLAMCloud);
    noiseRemoving(SLAMCloud);
    boxAdjustment(SLAMCloud);
    [S,R,T] = GeometricalTransformation(SLAMCloud, unitaryCarCloud);
    cleanCloud = linearTransformation(S,R,T,SLAMCloud);
}

```

FIGURE 10 Pseudo-code of the car detection process.

6 | ANALYSIS OF THE RESULTS

In this section we analyze the performance of the just presented methodology. Particular emphasis is made to the computational cost and real-time compliance of the on-line computations. A supplemental video showing both the process and the result of the proposed methodology can be downloaded from <https://youtu.be/Azm4E3Y7jS4>.

6.1 | Computer vision

The first part of the on-line procedure is the capture of the bodywork geometry through ORB-SLAM techniques. This algorithm runs in real time (i.e., it attains 30–60 frames per second) without any problem on a standard laptop. We employ the standard implementation in [50], whose code is freely available.

6.2 | Reconstruction of the geometry

Given the set of points captured in the process described in the previous section, the reconstruction of the geometry of the bodywork (i.e., the determination of the approximate geometry shown in Fig. 8) takes on average 0.30 seconds in a MacBook Pro 2016 computer equipped with 16 Gb RAM running an Intel Core i7 processor at 3.3 GHz.

6.3 | Embedding of the new model

After the reconstruction of the geometry, the new model or prototype should be embedded onto the manifold of shapes. This allows us to obtain the interpolation weights W_i , see Eq. (4). This process takes, on average, less than a second. Previous works have also reported how fast the LLE algorithm is, with some 500 different geometries running in about one second [38].

6.4 | Car identification

The procedure sketched in Fig. 10 runs in real time on a standard laptop.

6.5 | Error analysis

To test the accuracy of the LLE embedding onto the shape manifold, we deliberately kept five models out of the database. They were employed as ground truth, and their direct numerical simulation was compared to the result obtained by the interpolation

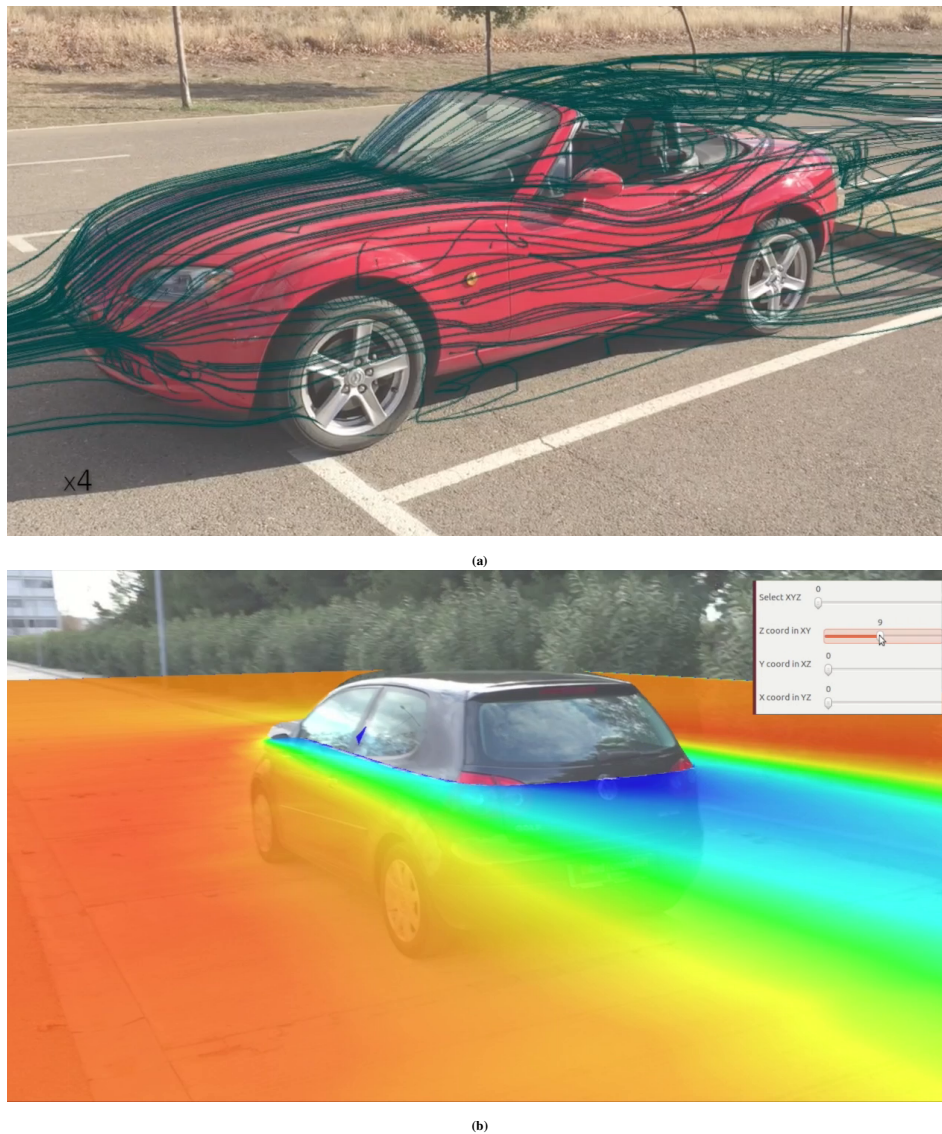


FIGURE 11 Final appearance of the augmented video (a) Streamtraces can be plotted on top of the reconstructed geometry. (b) Other possibilities, such as pressure or velocity contour plots, also exist.

of their velocity field, see Eq. (4). L_2 -norm errors ranged from 1.02% for the Hummer model to 0.56% for a classic car labelled as *sedan 4*, see Fig. 12 .

These results indicate that, for the scale of the grid employed in the characterization of geometry, many models provide very similar results. These velocity fields are, therefore, easy to interpolate. A highly refined mesh—which is not necessary for the type of applications we are envisaging here—would maybe report more important differences in the velocity field among different models.

7 | CONCLUSIONS

We have presented a tool able to capture and reconstruct, with the help of a commodity camera, the geometry of a three-dimensional object—we have studied mainly cars—and to perform simple but reasonably accurate CFD predictions of the flow around the object for interactive design. This technique employs state-of-the-art structure from motion techniques for the

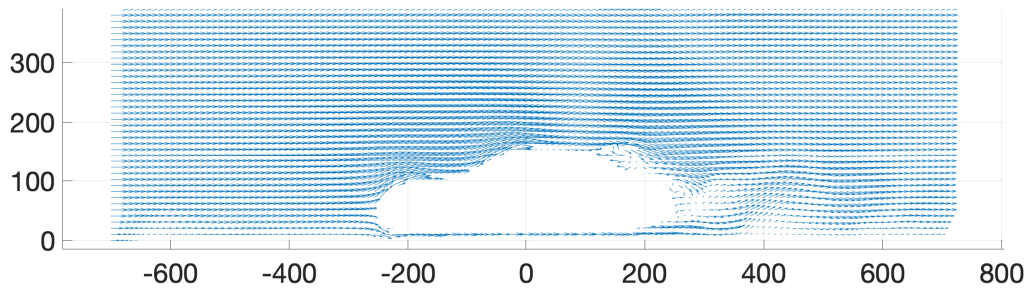


FIGURE 12 Interpolated velocity field for the *sedan 4* model. View along the symmetry plane.

geometrical reconstruction and manifold learning techniques to construct inexpensive predictions of the velocity and pressure fields around the object.

The advantage of such a platform is that what we obtain is actually an *appification* of CFD software. This tendency in our community is having tremendous commercial implications and is changing the way we see simulation. Here, the user needs only a smartphone to capture the geometry of the object to analyze, so as to obtain timely information about the physical consequences of his or her decisions. Artistic designers are thus invited to incorporate CFD to their workflow in a seamless way and with a minimum of knowledge in engineering—something that could be risky, nevertheless.

The combination of computer vision techniques, artificial intelligence (here, manifold learning) and augmented reality gives rise to a platform that changes the way we interact with simulation software. At the price of a little loss in accuracy with respect to existing, high-fidelity software, we obtain real-time responses to our query in a very short period of time, thus enabling decision making and an easy interplay between designers and engineers.

References

- [1] Bertram Anna, Othmer Carsten, Zimmermann Ralf. Towards Real-time Vehicle Aerodynamic Design via Multi-fidelity Data-driven Reduced Order Modeling. In: ; 2018.
- [2] Szeliski Richard. *Computer vision: algorithms and applications*. Springer International Publishing; 2010.
- [3] Hartley Richard, Zisserman Andrew. *Multiple View Geometry in computer vision*. Cambridge University Press; 2003.
- [4] Badías Alberto, Alfaro Icíar, González David, Chinesta Francisco, Cueto Elías. Reduced order modeling for physically-based augmented reality. *Computer Methods in Applied Mechanics and Engineering*. 2018;341:53 - 70.
- [5] Harwood Adrian RG, Wenisch Petra, Revell Alistair J. A REAL-TIME MODELLING AND SIMULATION PLATFORM FOR VIRTUAL ENGINEERING DESIGN AND ANALYSIS. In: ; 2018.
- [6] Mur-Artal R., Montiel J. M. M., Tardós J. D.. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. *IEEE Transactions on Robotics*. 2015;31(5):1552-3098.
- [7] Lassila Toni, Rozza Gianluigi. Parametric free-form shape design with PDE models and reduced basis method. *Computer Methods in Applied Mechanics and Engineering*. 2010;199(23):1583–1592.
- [8] Umetani Nobuyuki, Bickel Bernd. Learning Three-dimensional Flow for Interactive Aerodynamic Design. *ACM Trans. Graph.*. 2018;37(4):89:1–89:10.
- [9] Ghnatios Ch., Masson F., Huerta A., Leygue A., Cueto E., Chinesta F.. Proper Generalized Decomposition based dynamic data-driven control of thermal processes. *Computer Methods in Applied Mechanics and Engineering*. 2012;213-216(0):29 - 41.

- [10] Kirchdoerfer T., Ortiz M.. Data-driven computational mechanics. *Computer Methods in Applied Mechanics and Engineering*. 2016;304:81 - 101.
- [11] Ibañez Ruben, Borzacchiello Domenico, Aguado Jose Vicente, et al. Data-driven non-linear elasticity: constitutive manifold construction and problem discretization. *Computational Mechanics*. 2017;60(5):813–826.
- [12] Lam Remi R, Horesh Lior, Avron Haim, Willcox Karen E. Should You Derive, Or Let the Data Drive? An Optimization Framework for Hybrid First-Principles Data-Driven Modeling. *arXiv preprint arXiv:1711.04374*. 2017;.
- [13] Peherstorfer Benjamin, Willcox Karen. Dynamic data-driven reduced-order models. *Computer Methods in Applied Mechanics and Engineering*. 2015;291:21 - 41.
- [14] Latorre Marcos, Montáns Francisco Javier. What-You-Prescribe-Is-What-You-Get orthotropic hyperelasticity. *Computational Mechanics*. 2014;53(6):1279–1298.
- [15] Lee J.A., Verleysen M.. *Nonlinear Dimensionality Reduction*. Springer Verlag; 2007.
- [16] Díez Pedro, Zlotnik Sergio, Huerta Antonio. Generalized parametric solutions in Stokes flow. *Computer Methods in Applied Mechanics and Engineering*. 2017;326:223–240.
- [17] Manzoni Andrea, Quarteroni Alfio, Rozza Gianluigi. Model reduction techniques for fast blood flow simulation in parametrized geometries. *International Journal for Numerical Methods in Biomedical Engineering*. 2012;28(6-7):604–625.
- [18] Rozza G., Huynh D.B.P., Patera A.T.. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations – application to transport and continuum mechanics. *Archives of Computational Methods in Engineering*. 2008;15/3:229-275.
- [19] Manzoni A., Quarteroni A., Rozza G.. Computational reduction for parametrized PDEs: Strategies and applications. *Milan Journal of Mathematics*. 2012;80:283-309.
- [20] Manzoni Andrea, Quarteroni Alfio, Rozza Gianluigi. Shape optimization for viscous flows by reduced basis methods and free-form deformation. *International Journal for Numerical Methods in Fluids*. 2012;70(5):646–670.
- [21] Amsallem D., Cortial J., Farhat C.. Towards real-time CFD-based aeroelastic computations using a database of reduced-order information. *AIAA Journal*. 2010;48:2029-2037.
- [22] Amsallem D., Farhat Ch.. An Interpolation Method for Adapting Reduced-Order Models and Application to Aeroelasticity. *AIAA Journal*. 2008;46:1803-1813.
- [23] Oliehoek Frans A.. Interactive Learning and Decision Making: Foundations, Insights and Challenges. In: :5703–5708 International Joint Conferences on Artificial Intelligence Organization; 2018.
- [24] Chinesta Francisco, Cueto Elias. *PGD-Based Modeling of Materials, Structures and Processes*. Springer International Publishing Switzerland; 2014.
- [25] Gonzalez D., Masson F., Poulhaon F., Cueto E., Chinesta F.. Proper Generalized Decomposition based Dynamic Data Driven Inverse Identification. *Mathematics and Computers in Simulation*. 2012;82:1677-1695.
- [26] Kutz J Nathan. Deep learning in fluid dynamics. *Journal of Fluid Mechanics*. 2017;814:1–4.
- [27] Ghosh Pallab. AAAS: Machine learning 'causing science crisis' Published online, <https://www.bbc.com/news/science-environment-47267081> 2019.
- [28] Chang Michael B., Ullman Tomer, Torralba Antonio, Tenenbaum Joshua B.. A Compositional Object-Based Approach to Learning Physical Dynamics. *CoRR*. 2016;abs/1612.00341.
- [29] Bright Ido, Lin Guang, Kutz J Nathan. Compressive sensing based machine learning strategy for characterizing the flow around a cylinder with limited pressure measurements. *Physics of Fluids*. 2013;25(12):127102.

- [30] Jeong SoHyeon, Solenthaler Barbara, Pollefeys Marc, Gross Markus, others . Data-driven fluid simulations using regression forests. *ACM Transactions on Graphics (TOG)*. 2015;34(6):199.
- [31] Swischuk Renee, Mainini Laura, Peherstorfer Benjamin, Willcox Karen. Projection-based model reduction: Formulations for physics-based machine learning. *Computers & Fluids*. 2018;.
- [32] Raissi Maziar, Perdikaris Paris, Karniadakis George Em. Physics Informed Deep Learning (Part I): Data-driven solutions of nonlinear partial differential equations. *arXiv preprint arXiv:1711.10561*. 2017;.
- [33] González D., Chinesta F., Cueto E.. Thermodynamically consistent data-driven computational mechanics. *Continuum Mechanics and Thermodynamics*. 2018;DOI: 10.1007/s00161-018-0677-z.
- [34] Roweis Sam T., Saul Lawrence K.. Nonlinear Dimensionality Reduction by Locally Linear Embedding. *Science*. 2000;290(5500):2323-2326.
- [35] Rublee E., Rabaud V., Konolige K., Bradski G.. ORB: An efficient alternative to SIFT or SURF. In: :2564-2571; 2011.
- [36] Ballarin Francesco, Manzoni Andrea, Rozza Gianluigi, Salsa Sandro. Shape optimization by Free-Form Deformation: existence results and numerical solution for Stokes flows. *Journal of Scientific Computing*. 2014;60(3):537–563.
- [37] Quarteroni Alfio, Rozza Gianluigi. Optimal control and shape optimization of aorto-coronary bypass anastomoses. *Mathematical Models and Methods in Applied Sciences*. 2003;13(12):1801–1823.
- [38] González David, Cueto Elías, Chinesta Francisco. Computational Patient Avatars for Surgery Planning. *Annals of Biomedical Engineering*. 2015;44(1):35–45.
- [39] Lassila Toni, Manzoni Andrea, Quarteroni Alfio, Rozza Gianluigi. A reduced computational and geometrical framework for inverse problems in hemodynamics. *International journal for numerical methods in biomedical engineering*. 2013;29(7):741–776.
- [40] Zimmer Veronika A., Lekadir Karim, Hoogendoorn Corné, Frangi Alejandro F., Piella Gemma. A framework for optimal kernel-based manifold embedding of medical image data. *Computerized Medical Imaging and Graphics*. 2015;41:93 - 107. Machine Learning in Medical Imaging.
- [41] Agudo A., Moreno-Noguer F., Calvo B., Montiel J. M. M.. Sequential Non-Rigid Structure from Motion Using Physical Priors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2016;38(5):979-994.
- [42] Civera J., Bueno D. R., Davison A. J., Montiel J. M. M.. Camera self-calibration for sequential Bayesian structure from motion. In: ; 2009.
- [43] Bradski Gary, Kaehler Adrian. *Learning OpenCV: Computer vision with the OpenCV library*. " O'Reilly Media, Inc."; 2008.
- [44] Edelsbrunner Herbert, Mücke Ernst P. Three Dimensional Alpha Shapes. *ACM Transactions on Graphics*. 1994;13:43-72.
- [45] Schuon Sebastian, Durković Marko, Diepold Klaus, Scheuerle Jürgen, Markward Stefan. Truly incremental locally linear embedding. In: ; 2008.
- [46] Mahmoud Nader, Collins Toby, Hostettler Alexandre, Soler Luc, Doignon Christophe, Montiel Jose Maria Martinez. Live Tracking and Dense Reconstruction for Handheld Monocular Endoscopy. *IEEE transactions on medical imaging*. 2019;38(1):79–89.
- [47] Mousavian Arsalan, Anguelov Dragomir, Flynn John, Kosecka Jana. 3d bounding box estimation using deep learning and geometry. In: :7074–7082; 2017.
- [48] Hödlmoser Michael, Micusik Branislav, Liu Ming-Yu, Pollefeys Marc, Kampel Martin. Classification and pose estimation of vehicles in videos by 3d modeling within discrete-continuous optimization. In: :198–205IEEE; 2012.
- [49] Shreiner Dave, The Khronos OpenGL ARB Working Group . *OpenGL Programming Guide: The Official Guide to Learning OpenGL, Versions 3.0 and 3.1*. Addison-Wesley Professional; 7th ed.2009.

- [50] Mur-Artal Raul, Tardós Juan D. Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras. *IEEE Transactions on Robotics*. 2017;33(5):1255–1262.

