Neural Implicit Surface Evolution

Abstract

This work investigates the use of smooth neural networks for modeling dynamic variations of implicit surfaces under the level set equation (LSE). For this, it extends the representation of neural implicit surfaces to the space-time $\mathbb{R}^3 \times \mathbb{R}$, which opens up mechanisms for continuous geometric transformations. Examples include evolving an initial surface towards general vector fields, smoothing and sharpening using the mean curvature equation, and interpolations of initial conditions.

The network training considers two constraints. A data term is responsible for fitting the initial condition to the corresponding time instant, usually $\mathbb{R}^3 \times \{0\}$. Then, a LSE term forces the network to approximate the underlying geometric evolution given by the LSE, without any supervision. The network can also be initialized based on previously trained initial conditions, resulting in faster convergence compared to the standard approach.

1. Introduction

A neural implicit function $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ is a smooth neural network that represents an implicit function. Since $g$ is smooth, objects from the differential geometry of its regular level sets can be used in closed form [37, 27].

This work investigates the extension of the domain of neural implicit functions to the space-time $\mathbb{R}^3 \times \mathbb{R}$, encoding the evolution of the function $g$ as a higher-dimensional function $f : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}$. The resulting animation is governed by a PDE, the level set equation (LSE) $\frac{\partial f}{\partial t} = v \| \nabla f \|$, which encodes the propagation of the level sets $S_t$ of $f(\cdot, t)$ towards their normals with speed $v$. The choice of the function $v$ depends on the underlying geometric model. LSE is an important tool for geometry processing applications.

We propose to use a neural network $f : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}$ to represent the above level set function. For this, we train $f$ to learn the evolution $S_t$ of an initial surface $S$ under a given LSE. Accordingly, we add the constraint $f = g$ on $\mathbb{R}^3 \times \{0\}$, to the LSE. If $g$ is the signed distance function (SDF) of $S$, then a solution $f$ of this problem encodes an animation of $S$.

Our method is the first neural approach in geometry processing, that does not consider numerical approximations of the LSE solution during sampling and does not discretize the LSE in the loss function.

Our strategy has two steps to train $f$. First, a sample $\{p_i, g(p_i)\}$ enables us to train the initial condition $f = g$ on $\mathbb{R}^3 \times \{0\}$. Second, an LSE constraint is used to fit $f$ into a solution of the LSE in $\mathbb{R}^3 \times \mathbb{R}$. This term does not need any supervision, i.e. it does not consider samples of $f$. This constraint only uses samples in the form $(p_i, t_i)$. The only requirement of our method is that $f$ must be smooth. Given this, our main contributions can be summarized as follows:

- Extension of neural implicit surfaces to space-time without the use of numerical/discrete approximations. Encoding the space-time coordinates as input of the network allows us to represent the whole implicit animation in a single network;

- Development of a neural framework to fit solutions of an LSE using only its analytical expression (Sec 4). Moreover, the network training considers only the initial condition of the LSE problem as data;

- The method is flexible enough to be used in a variety of applications, such as surface motion by vector fields, smoothing, sharpening, and interpolation (Sec 6);

- We also propose a novel network initialization based on previously trained initial conditions (Sec 4.4).

2. Related Works

Many problems, such as shape correspondence, topology changes, and animation of deformable objects, can be posed using implicit surfaces [17, 34, 3, 4, 14, 13]. Studying their shape properties through differential geometry leads to a framework for intrinsic operations. An example is smoothing a surface using the mean curvature equation, an important PDE in geometry processing [6, 5, 15, 13]. Problems in this topic rely on computing derivativess – a hard task when dealing with meshes [7, 11, 12]. A practical neural implicit approach would allow us to compute such objects in closed form, and it is the objective of this work.
Several works showed that modeling surfaces as level sets of neural networks result in a compact representation [30, 26, 18, 33, 27, 10]. Most of them fit a network into data. In the SDF case, a regularizer term forces the network to satisfy the Eikonal equation. The robustness of those approaches is our motivation to study the evolution of neural implicit surfaces using the level set equation [29, 28], a PDE widely used in geometry processing [13, 2, 21, 32, 24, 36].

NIE [37] and NIE [25] are recent neural approaches to evolve implicit surfaces. Both store the evolution in a sequence of networks, each representing a time step, as opposed to our method, which encodes it in a single network. Specifically, they use a network \( g_\phi \) to fit an initial function \( g \). Then, \( \phi \) is updated at each time step creating a sequence of networks \( g_\phi : \mathbb{R}^3 \to \mathbb{R} \). This is similar to the Runge-Kutta methods but, instead of fitting the solution to a grid, they use a network. Thus, for them to evaluate at intermediate times, they have to retrain the networks. In contrast, our method does not discretize time, learning the solution in a continuous interval using a single network with a domain in \( \mathbb{R}^3 \times \mathbb{R} \).

NIE is the only approach that evolves \( g_\phi \) using the mean curvature equation \( \frac{\partial^2 \phi}{\partial t^2} = \| \nabla f \| \kappa \). However, this evolution uses a discretization of the partial derivative \( \frac{\partial^2 \phi}{\partial t^2} \). To update \( \phi \), NIE uses a finite difference scheme to approximate the discrete solutions. To compute the mean curvature \( \kappa \), NIE extracts the level sets using marching cubes and employs the cotangent Laplacian, which is problematic since it depends on approximating the level sets by meshes using marching cubes. Also, this operator does not preserve the Laplacian natural properties – the no free lunch scenario [35]. In contrast, our approach uses the network high-order derivatives to evaluate the LSE analytically. For example, we compute the curvature using the divergence of \( \frac{\nabla^2 f}{\| \nabla f \|} \). To evaluate \( \frac{\nabla^2 f}{\| \nabla f \|} \) and \( \nabla f \) we simply use automatic differentiation. Moreover, NIE cannot consider multiple initial conditions as our approach does.

Recently, there has been a growing interest by the physical simulation community in solving PDEs using neural networks. Physically informed neural networks (PINNs) are established approaches in this context. Unlike our proposal, this method [16] relies on measurements of the PDE solution at intermediate times. PINNs are extensively evaluated in surveys [20, 8]. Karniadakis et al. [20] reviews inverse problems, which try to infer the PDE parameters based on supervised data. This context differs from ours since we do not address the inverse problem, they also do not consider implicit surface evolution using the LSE, and we do not rely on supervised data of the evolution.

Cuomo et al. [8] surveys a broader range of problems based on the PDE type. We focus on the evolution of implicit surfaces using the LSE, which is a geometric time-dependent PDE not explored in their review [8, Sec 3.2.2]. Thus, to the best of our knowledge, there is no PINN-based approach to solve the LSE for implicit surface evolution.

Our proposal seeks to bridge this gap by leveraging the representation capacity of (sinusoidal coord-based) networks to solve the above geometric problem without any measurements of the PDE solution at intermediate times; i.e., we need supervision only on the initial surfaces. This can potentially enable new applications in computer graphics, computer-aided design, and computational geometry.

Regarding the interpolation problem, Liu et al. [23] propose the Lipschitz MLP, which regularizes a neural network by penalizing the upper bound of its Lipschitz constant. We use a specific PDE to interpolate SDFs, resulting in smoother and more natural transitions between shapes (see Sec 6.3). Moreover, our method manages to use smaller architectures by considering sinusoidal MLPs.

3. Background and conceptualization

3.1. Implicit surfaces

The level set \( g^{-1}(0) \) of a smooth function \( g : \mathbb{R}^3 \to \mathbb{R} \) is a (regular) surface if \( \nabla g \neq 0 \) in \( g^{-1}(0) \). Conversely, given a surface \( S \), there is a function \( g \) having it as its zero-level set. Thus, \( g \) may be reconstructed from a sample of \( S \). For this, we parametrize \( g \) using a neural network. SIREN [33] and IGR [18] are examples of such networks.

To compute the parameters of \( g \) such that \( g^{-1}(0) \approx S \), it is common to consider the Eikonal problem:

\[
\| \nabla g \| = 1 \text{ subject to } g = 0 \text{ on } S. \tag{1}
\]

Which asks for \( g \) to be the SDF of a set containing \( S \). We can derive from Eq (1) that \( \langle \nabla g, N \rangle = 1 \) on \( S \), which implies that \( \nabla g \) must be aligned with the normals of \( S \).

We are interested in using neural networks to evolve \( S \). The level sets of \( g \) could be used, but they do not allow intersections between surfaces at different instants. To avoid this, we extend the domain of the implicit function to \( \mathbb{R}^3 \times \mathbb{R} \), where the parameter \( t \in \mathbb{R} \) controls the animation.

3.2. Evolving implicit surfaces

We use a function \( f : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R} \) to define the above extension. For evolving the level sets of \( g \) we require \( f = g \) in \( \mathbb{R}^3 \times \{0\} \) and the resulting evolution is given by

\[
S_t = f^{-1}_t(0) = \{ p \in \mathbb{R}^3 | f(p, t) = 0 \} \tag{2}
\]

We assume that \( f \) is negative in the interior of \( S_t \) and its normal vectors are given by \( N = \frac{\nabla f}{\| \nabla f \|} \). Furthermore, \( \nabla f \) denotes the gradient of \( f \) with respect to the space \( \mathbb{R}^3 \).

The time \( t \) allows continuous navigation in \( S_t \) and represents a transformation \( S_0 \to S_t \) of the initial surface \( S_0 \).
Moreover, for \( t_0, t_1 \in \mathbb{R} \), the function \( f \) provides a smooth deformation between \( S_{t_0} \) and \( S_{t_1} \). These surfaces can contain singularities as their topologies may change over time. Hart [19] has studied this phenomenon using Morse theory.

Such an Eulerian approach is usually explored in geometry processing by storing \( f \) on a 4D grid [13]. Here, we take an analytical formulation by parametrizing \( f \) by a (coordinate-based) network. Such an approach has several advantages. First, automatic differentiation provides us the analytical derivatives of \( f \), which may be used at the loss function. We can also compute the normals and curvature measures of \( S_t \) in closed form [27]. Additionally, neural networks are compact representations for implicit functions, guaranteed by the universal approximation theorem [9]. Storing \( f \) with precision in a 4D grid could be an unfeasible task.

### 3.3. Level set equation

Encoding the surface evolution \( S_t \) by the time-dependent function \( f \) results in PDE – the level set equation (LSE). We describe it below.

Let \( g : \mathbb{R}^3 \to \mathbb{R} \) be the SDF of the initial surface \( S \). A function \( f : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R} \) encodes the evolution \( S_t \) if \( f = g \) on \( \mathbb{R}^3 \times \{0\} \), and for each point \( p \), the function \( f \) is constant along its path \( \alpha(t) \), i.e., \( f(\alpha(t), t) = c \) iff \( g(p) = c \). Thus, deriving the function \( f(\alpha(t), t) \) we obtain

\[
\frac{\partial f}{\partial t}(\alpha(t), t) + \langle \nabla f(\alpha(t), t), \alpha'(t) \rangle = 0.
\]

The derivative \( \alpha'(t) \) is a vector field along the path \( \alpha(t) \).

As Eq (3) holds for each point \( p \), we can drop its path \( \alpha \) and use only \( \alpha' \), which can be seen as a time-dependent vector field \( V(p, t) = \alpha'(t) \). Then, the function \( f \) encoding the animation \( S_t \) is a solution of the LSE:

\[
\begin{cases}
\frac{\partial f}{\partial t} + \langle \nabla f, V \rangle = 0 & \text{in } \mathbb{R}^3 \times (a, b), \\
f = g & \text{on } \mathbb{R}^3 \times \{t = 0\}.
\end{cases}
\]

The time interval \( (a, b) \) contains \( t = 0 \) and controls the evolution \( S_t \). A solution \( f \) of Eq (4) implicitly encodes the integration of \( V \). Thus, defining a family of vector fields is a way to animate a given surface using the LSE.

Observe that in the same sense that we considered a neural implicit function as a solution of Eq (1), we assume that the function \( f \) is a solution of \( \frac{\partial f}{\partial t} + \langle \nabla f, V \rangle = 0 \) subject to \( f = g \) on \( \mathbb{R}^3 \times \{0\} \). Sec 5 presents examples of LSEs.

### 4. Method

We propose representing a surface evolution by a (coordinate-based) neural network \( f_\theta : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R} \). Given an LSE with initial conditions, this section defines a machine learning framework, consisting of a loss functional, sampling strategies, and a network initialization, to train \( f_\theta \) to approximate a solution to the LSE problem.

Let \( g_i : \mathbb{R}^3 \to \mathbb{R} \) be the SDFs of \( n \) surfaces \( S_i \). We train \( f_\theta \) by forcing it to approximate a solution a neural LSE:

\[
\begin{align*}
\mathcal{F} := & \frac{\partial f_\theta}{\partial t} + \langle \nabla f_\theta, V \rangle = 0 \quad \text{in } \mathbb{R}^3 \times (a, b), \\
f_\theta = g_i & \quad \text{on } \mathbb{R}^3 \times \{t_i\}.
\end{align*}
\]

We employed the notation \( \mathcal{F} \) to represent the LSE for brevity. The untrained network \( f_\theta \) must encode the movement ruled by the vector field \( V \). \((a, b)\) can be used to control the resulting neural animation \( S_t \) of \( S \).

### 4.1. Loss functional

We use Eq (5) to define a loss function to train \( f_\theta \).

\[
\mathcal{L}(\theta) = \int_{\mathbb{R}^3 \times (a, b)} |\mathcal{F}| \, dp + \sum_{i=1}^{n} \int_{\mathbb{R}^3 \times \{t_i\}} |f_\theta - g_i| \, dp.
\]

The LSE constraint \( \mathcal{L}_\text{LSE} \) forces the network \( f_\theta \) to satisfy \( \mathcal{F} = 0 \) and works as a regularization of \( f_\theta \) that requires it to follow the underlying phenomenon. The data constraint \( \mathcal{L}_\text{data} \) asks for \( f_\theta \) to satisfies \( f_\theta = g_i \) on \( \mathbb{R}^3 \times \{t_i\} \).

### 4.2. Sampling

To approximate a solution \( f_\theta \) of Eq (5), we seek a minimum \( \theta \) of the loss function \( \mathcal{L} \) using the stochastic gradient descent. For this, we enforce \( \mathcal{L} = \mathcal{L}_\text{LSE} + \mathcal{L}_\text{data} \) with a sampling in \( \mathbb{R}^3 \times (a, b) \) and another in \( \bigcup_{i=1}^{n} \mathbb{R}^3 \times \{t_i\} \).

#### Sampling space-time points

During training, we sample minibatches of \( l_i \in \mathbb{N} \) space-time points \( (p_j, t_j) \in \mathbb{R}^3 \times (a, b) \) randomly. Then, \( \mathcal{L}_\text{LSE} \) is enforced in \( (p_j, t_j) \), yielding the approximation:

\[
\hat{\mathcal{L}}_\text{LSE}(\theta) = \frac{1}{l_i} \sum_{j=1}^{l_i} |\mathcal{F}(p_j, t_j)|
\]

Observe that the LSE constraint \( \mathcal{L}_\text{LSE} \) does not need any data supervision.

#### Sampling initial conditions

The data constraint \( \mathcal{L}_\text{data} \) forces \( f_\theta \) to fit the input dataset, which we consider to be the SDFs \( g_i \) of \( n \) surfaces \( S_i \).

We use Eq (1) to define \( \mathcal{L}_\text{data} = \sum_i \mathcal{L}_i \), with \( \mathcal{L}_i \) managing the restrictions \( g_i = f_\theta \) on \( \mathbb{R}^3 \times \{t_i\} \).

\[
\mathcal{L}_i = \left[ 1 - \langle \nabla f_\theta, N_i \rangle \right] \, dp + \left[ \int f_\theta - g_i \, dp + \int [1 - \langle \nabla f_\theta, N_i \rangle \, ds_i \right]
\]

Where \( \mathcal{L}_\text{Dirichlet} \) asks for \( f_\theta \) to fit \( g_i \) at time \( t_i \), \( \mathcal{L}_\text{Neumann} \) requires the alignment between \( \nabla f_\theta \) and the normals of \( S_i \),...
and \( \mathcal{L}_{\text{Eikonal}} \) is the Eikonal regularization. During training, these constraints are discretized, as in the PDE constraint case. Then, we sample minibatches with \( l_2 \) on-surface points \( (g_i = 0) \) and \( l_3 \) off-surface points \( (g_i \neq 0) \). Observe that only the initial conditions \( \{g_i\} \) are used in \( \mathcal{L}_{\text{data}} \).

In practice, we use two kinds of initial conditions. First, the neural networks \( g_i \) fit the SDFs of \( S_i \), resulting in faster training since, for each point \( p_i \), the values \( g_i(p_i) \) and \( \nabla g_i \) given by the evaluation of \( g_i \) and its derivative at \( p_i \). In this case, we avoid using \( \mathcal{L}_{\text{Eikonal}} \) since \( g_i \) are already trained to satisfy the Eikonal equation. Second, we consider point clouds \( \{p_i, N_i\}_i \) sampled from \( S_i \), where we have to approximate the SDF of \( S_i \) [27]. In both cases, we include the constraint \( \mathcal{L}_{\text{Neumann}} \) that forces a normal alignment at \( t = 0 \) regularizing the orientation near the zero-level set.

During training, we sample minibatches of size \( l_1 + l_2 + l_3 \) to feed \( \mathcal{L} \); \( l_i \) are the numbers of space-time, on-surface, and off-surface points. The experiments shown good results using that \( l_1, l_2, l_3 \) have 50\%, 25\%, 25\% of the minibatch size. The supplementary materials give experiments varying \( l_i \).

### 4.3. Neural network architecture

We consider the neural network to be a sinusoidal MLP
\[
f_\theta(p) = W_{d+1} \circ f_d \circ \cdots \circ f_1(p) + b_{d+1},
\]
with \( d \) hidden layers \( f_i(p_l) = \sin(W_i p_l + b_i) \), where \( W_i \in \mathbb{R}^{N_i+1 \times N_i} \) are the weight matrices, and \( b_i \in \mathbb{R}^{N_i+1} \) are the biases. The sine is applied at each coordinate of \( W_i p_l + b_i \), \( \theta \) consists of the union of the coefficients of \( W_i \) and \( b_i \). The integer \( d \) is the depth of \( f_\theta \) and the dimensions \( N_i \) are the layers widths.

The network \( f_\theta \) is smooth and we can compute its derivatives using automatic differentiation. Therefore, we train \( f_\theta \) using the loss function \( \mathcal{L} \) in Eq (6).

### 4.4. Network initialization

We introduce a novel initialization of \( f_\theta : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R} \) based on a previously trained network \( g_\phi : \mathbb{R}^3 \to \mathbb{R} \). This initialization of \( \theta \) using \( \phi \) results in faster training compared with the standard definitions [33] (see Sec 6.4).

Assume that the training of \( f_\theta \) is subject to \( f_\theta = g_\phi \) on \( \mathbb{R}^3 \times \{0\} \). Then we define \( \theta \) in terms of \( \phi \) such that \( f_\theta(p, t) = g_\phi(p) \) for all \( t \), that is, \( f_\theta \) will be constant and equal to \( g_\phi \) over time. This allows us to start the training of \( f_\theta \) to fit a solution of an underlying LSE problem with \( f_\theta \) already satisfying the initial condition.

For this, we suppose that \( f_\phi \) is wider than \( g_\phi \) and that their depths are equal to \( d \). Specifically, let \( B_i \in \mathbb{R}^{N_i+1 \times N_i} \), \( b_i \in \mathbb{R}^{N_i+1} \) be the trained weight matrices and biases of \( g_\phi \), and \( A_i \in \mathbb{R}^{M_i+1 \times M_i} \), \( a_i \in \mathbb{R}^{M_i+1} \) be the untrained weight matrices and biases of \( f_\phi \). Since \( f_\phi \) is wider than \( g_\phi \), i.e. \( N_i \leq M_i \), we can define \( A_i, b_i \) using
\[
A_1 = \begin{pmatrix} B_1 & 0 \\ F_1 & 0 \end{pmatrix}, \quad A_i = \begin{pmatrix} B_i & 0 \\ 0 & 0 \end{pmatrix} \quad \text{for } i = 2, \ldots, d,
A_{d+1} = \begin{pmatrix} B_{d+1} & L \end{pmatrix}, \quad a_i = \begin{pmatrix} b_i \\ 0 \end{pmatrix} \quad \text{for } i = 1, \ldots, d + 1.
\]

Thus, \( f_\theta(p, t) = g(p) \) for all \( t \), see supp. material for the details. \( F_p, F_t \) project the input \((p, t)\) in the dictionary \( \sin(F_p p + F_t t) \), and are initialized using the standard approach. Note that these sines are not used in the first training step, but as it advances, the new hidden weights combine them improving the training (see Sec 6.4).

### 5. Examples

Here, we present examples of neural implicit evolution using LSE. Sec 5.1 shows examples using time-independent vector fields. Sec 5.2 considers the mean curvature equation, which is intrinsically related to the surface and results in smoothing/sharpening applications for implicit neural surfaces. Sec 5.3 investigates interpolations between implicit neural surfaces using an LSE.

Recently, these problems have been addressed by different neural methods [25, 37, 23]. Comparisons are made in Sec 6. Hereafter, we give each application’s conceptualization and corresponding loss function.

#### 5.1. Time-independent vector fields

Moving a surface \( S \) towards a vector field \( V : \mathbb{R}^3 \to \mathbb{R}^3 \) results in a simple LSE. Specifically, let \( g \) be the SDF of \( S \). Since \( V \) does not change over time, it may be defined and customized beforehand. For example, sources, sinks, saddles, and constant vectors may be used to generate vector fields based on specific applications.

We train a neural network \( f_\theta : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R} \) to implicitly encode the evolution of \( S \) by \( V \) using the resulting LSE:
\[
\begin{align*}
\frac{\partial f_\theta}{\partial t} - v \| \nabla f_\theta \| &= 0 \quad \text{in } \mathbb{R}^3 \times (a, b), \\
f_\theta &= g \quad \text{on } \mathbb{R}^3 \times \{0\}.
\end{align*}
\]

Here, \( v \) denotes the size of the normal component of \( V \), that is, \( v(p, t) = \langle V(p), N(p) \rangle \). The minus in Eq (7) is because we need the inverse of the resulting flow to compose with \( g \).

Sec 6.1 gives two experiments using time-independent vector fields as a proof of concept. Other examples are presented in the video supplementary material.

#### 5.2. The mean curvature equation

The mean curvature equation evolves the level sets with velocity given by the negative of their mean curvature, resulting in a smoothing along the time [6, 1].

Let \( V(p, t) = -\kappa(p, t)N(p, t) \) be the mean curvature vector, where \( N \) is the normal field of the level sets and \( \kappa = \text{div } N \) is the mean curvature; \( \text{div} \) is the divergence operator. Replacing \( V \) in Eq (5) results in:
\[
\begin{align*}
\frac{\partial f}{\partial t} - \alpha \| \nabla f \| \kappa &= 0 \quad \text{in } \mathbb{R}^3 \times (a, b), \\
f &= g \quad \text{on } \mathbb{R}^3 \times \{t = 0\}.
\end{align*}
\]
Intuitively, the zero-level set moves toward the mean curvature vector $-\kappa N$, contracting regions with positive curvature and expanding regions with negative curvature. Thus, such procedure smooths (sharpen) the surface if $t > 0$ ($t < 0$). $\alpha$ controls the level set evolution, which has relations with minimal surfaces (Sec 1 of supp. material).

Analogously to Eq. (5), we define $LSE + L_{data}$ to fit a solution of Eq. (8) using $F := \frac{\partial f}{\partial t} - \alpha \| \nabla f \| \phi_0$. Sec 6.2 presents smoothing and sharpening using this technique.

5.3. Interpolation between implicit surfaces

Let $g_i$ be the SDFs of two surfaces $S_i$. We present a LSE approach to interpolate $g_i$. A vector field $V$ for Eq (5), such that its solution interpolates between $S_i$, has the form:

$$V(p, t) = -(g_2(p) - f(p, t)) \frac{\nabla f(t, p)}{\| \nabla f(t, p) \|}, \quad (9)$$

with $f(p, 0) = g_1(p)$. Note that the evolution towards $V$ forces each $c$-level set of $g_1$ to match the $c$-level set of $g_2$. The resulting LSE is given by substituting Eq (9) in Eq (5)

$$\begin{cases} \frac{\partial f}{\partial t} - \| \nabla f \| (g_2 - f) = 0 & \text{in } \mathbb{R}^3 \times \mathbb{R}, \\ f = g_i & \text{on } \mathbb{R}^3 \times \{t_i\}. \end{cases} \quad (10)$$

A solution $f$ of Eq (10) will locally inflate $S_1$ if inside $S_2$, and deflate it if outside so that $S_1$ will always try to fit into $S_2$ [13]. Again, we define a loss function $LSE + L_{data}$ to fit a solution of Eq (10) using $F := \frac{\partial f}{\partial t} - \| \nabla f \| (g_2 - f)$.

Theoretically, we could use the mean curvature equation to minimize deformations along the interpolation. This LSE has the property of minimizing area distortions of the resulting evolution; see Sec 1 of the supplementary material.

6. Experiments

Here, we present the experiments of the examples given in Sec 5. See the supplementary material for an ablation study of the training, sampling, and initialization.

6.1. Deformation driven by vector fields

We use the definitions in Sec 5.1 to train an animation based on vector fields spatially related to the initial surface. First, consider the twist $V(x, y, z) = y(-z, x, 0)$ of $\mathbb{R}^3$ along the $y$-axis. Substituting it in Eq (5) results in a level set equation, which we use to derive a loss function.

Let $g$ be the SDF of the Armadillo and $f_0$ be a network with 2 hidden layers $f : \mathbb{R}^{256} \rightarrow \mathbb{R}^{256}$. We trained $f_0$ during 48000 epochs using minibatches of 25000 on-surface points ($g = 0$), 25000 off-surface points ($g \neq 0$), and 8000 points in $\mathbb{R}^3 \times [-1, 1]$. Fig 1 shows 3 reconstructions of the zero-level sets of $f_0$ at times $t_i = 0, 0.25, 0.5$.

Although we do not provide data in $\mathbb{R}^3 \times \{t \neq 0\}$, the solution is well approximated (see Fig 1). The vertical axis is the $y$-axis, and the origin of $\mathbb{R}^3$ is at the ground.

For the next experiment, let $g$ be the SDF of the Spot (Fig 2, center). Define $V = V_1 - V_2$ as the sum of a source $V_1$ and a sink $-V_2$, with $V_i(p) = e^{-p_i^2}$ at times $t_i = 0, 256$. A vector field that represents a twist of $\mathbb{R}^3$ along the $y$-axis.

The points $p_1$ and $p_2$ are the centers of Spot’s body and head. Again, we use $V$ to derive a loss function to train $f_0$. We parameterize $f_0$ with one hidden layer $f : \mathbb{R}^{128} \rightarrow \mathbb{R}^{128}$ and train it for 46000 epochs. As expected, it reduces the Spot’s head while it increases the body size, see Fig 2.

Table 1 shows that the above networks are close to satisfying the LSE problems. We compare $f_0(\cdot, 0)$ with the initial condition $g$ and measure how close $f_0$ is from satisfying $F = 0$. For this we use the following measures: 1) The absolute difference $|f_0(\cdot, 0) - g|$ in $\mathbb{R}^3 \times \{0\}$; 2) The evaluation of $f_0$ in $|F|$ in $\mathbb{R}^3 \times \mathbb{R}$. We use a sample of 1000 points in $\mathbb{R}^3 \times \{0\}$ and $\mathbb{R}^3 \times \mathbb{R}$, not included in the training process, to evaluate the mean/maximum of each measure.

<table>
<thead>
<tr>
<th>Vector field</th>
<th>on-surface constraint mean</th>
<th>on-surface constraint max</th>
<th>off-surface constraint mean</th>
<th>off-surface constraint max</th>
<th>PDE constraint mean</th>
<th>PDE constraint max</th>
</tr>
</thead>
<tbody>
<tr>
<td>twist</td>
<td>0.0008</td>
<td>0.0003</td>
<td>0.002</td>
<td>0.028</td>
<td>1e-5</td>
<td>0.0004</td>
</tr>
<tr>
<td>source-sink</td>
<td>0.0009</td>
<td>0.0005</td>
<td>0.001</td>
<td>0.015</td>
<td>2.7e-6</td>
<td>0.0005</td>
</tr>
</tbody>
</table>

Table 1. Comparisons between the ground truth initial conditions at $t = 0$ and the measures of how close the trained networks are to satisfy the underlying LSEs.
6.2. The mean curvature equation

The next set of experiments seek to solve the mean curvature equation. First, we consider simple initial conditions such as the cube and the dumbbell. Then we use the intrinsic properties of this LSE to provide smoothing/sharpening of detailed surfaces.

6.2.1 Simple initial conditions for validation

Let \( g \) be the SDF of the cube, and \( f : \mathbb{R}^3 \rightarrow \mathbb{R} \) be a network with 3 hidden layers \( f_i : \mathbb{R}^{256} \rightarrow \mathbb{R}^{256} \). We set \( \alpha = 0.1 \) and optimized \( f \) for 8000 epochs using the loss function resulting from Eq (8). We used an oriented point cloud of size 40000, sampled from the cube. During training, we consider minibatches of 5000 on-surface points, 5000 off-surface points, and 10000 in \( \mathbb{R}^3 \times [-1, 1] \).

Fig 3 shows the level sets of \( f \) at times \( t_i \) for \( i = 0, \ldots, 5 \). As expected, regions with positive mean curvature, such as the cube corners, contract. This LSE evolves the surface toward the normals times the negative of the mean curvature. Therefore, the cube will at some instant of time collapse to a point, but right before it will be very close to a sphere.

The dumbbell is a classical example. Let \( g \) be its SDF and \( f_\theta \) be a network with 2 hidden layers \( f_i : \mathbb{R}^{256} \rightarrow \mathbb{R}^{256} \). We set \( \alpha = 0.05 \) and sampled a point cloud of size 80000 from the dumbbell. The training took 2800 epochs using minibatches of 5000 on-surface points, 5000 off-surface points, and 10000 in \( \mathbb{R}^3 \times [-1, 1] \).

Fig 4 shows the level sets of \( f_\theta \) at times \( t_i = \frac{i}{10} \) for \( i = 0, \ldots, 7 \). As expected, since the neck region has higher mean curvature, it pinches off first creating two connected components. Later, each component collapses to a point, becoming small spheres right before that. The resulting flow has critical points in different instances of time.

6.2.2 Smoothing and sharpening

The mean curvature equation evolves the level sets by contracting (expanding) regions with positive (negative) mean curvature. As a consequence, its solution smooths (sharpens) the level sets when \( t > 0 \) (\( t < 0 \)).

Let \( g \) be the SDF of the Armadillo, and \( f_\theta \) be a network with 2 hidden layers \( f_i : \mathbb{R}^{256} \rightarrow \mathbb{R}^{256} \). We set \( \alpha = 0.001 \) in Eq (8). The network \( f_\theta \) was trained during 33000 epochs using an oriented point cloud of size 80000. During training, we used minibatches of 5000 on-surface points, 5000 off-surface points, and 10000 in \( \mathbb{R}^3 \times [-1, 1] \).

Fig 5 presents three reconstructions of the zero-level sets of \( f_\theta \) at times \( t = 0, 0.1, 0.2 \). As expected, the Armadillo surface was properly reconstructed at \( t = 0 \) and, as time progressed, it became smoother. Regions with positive mean curvature, such as the fingers, contracted.

For the sharpening we reconstruct the zero-level sets at \( t = 0, -0.1, -0.2 \) (see Fig 6). As expected, regions with positive curvature have expanded, resulting in an enhancement of the geometrical features of the surface.
Numerical evaluation

There is no available (ground-truth) analytical solutions of the mean curvature equation (Eq \((\ref{eq:mean_curvature})\)) for the above surfaces and finding them is not trivial. Nonetheless, we can quantitatively evaluate the network’s proximity to satisfying Eq \((\ref{eq:mean_curvature})\). This is presented in Table 2. We compare the trained networks at \(t = 0\) with the initial surfaces. We also measure how close the networks are to satisfying the LSE. We used a sample of 1000 points in \(\mathbb{R}^3 \times \{0\}\) and \(\mathbb{R}^3 \times \mathbb{R}\), not included in the training process, to evaluate the mean and maximum values of each measure.

<table>
<thead>
<tr>
<th>Model</th>
<th>on-surface constraint</th>
<th>off-surface constraint</th>
<th>PDE constraint</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>max</td>
<td>mean</td>
</tr>
<tr>
<td>Cube</td>
<td>0.0006</td>
<td>0.007</td>
<td>0.0013</td>
</tr>
<tr>
<td>Dumbbell</td>
<td>0.0003</td>
<td>0.002</td>
<td>0.0010</td>
</tr>
<tr>
<td>Armadillo</td>
<td>0.0008</td>
<td>0.004</td>
<td>0.0022</td>
</tr>
</tbody>
</table>

Table 2. Quantitative evaluation of our method in the problem of approximating solutions of the mean curvature equation.

Comparisons

We compare our technique with NFGP [37] and NIE [25] for smoothing and sharpening of neural implicit surfaces.

NFGP evolves a network \(g_\phi : \mathbb{R}^3 \rightarrow \mathbb{R}\) such that the level set of the resulting network \(g_\phi\) smooths/sharpens \(g_\phi^{-1}(0)\). The training optimizes, \((\kappa_\phi - \beta \kappa_\phi)^2\), the difference between the mean curvatures of the level sets of \(g_\phi\) and \(g_\phi\). Then, using \(\beta < 1\) (\(\beta > 1\)), it would force a smoothing (sharpening) of the initial surface. However, NFGP trains a network \(g_\phi\) for each \(\beta\), thus it cannot represent a continuous evolution over time. In contrast, our framework directly evolves over time using a single network. Although the NFGP approach does not use the mean curvature equation model, we can still perform a qualitative analysis as a means of comparison, since a numerical analysis is not feasible. Fig 7 shows this comparison for sharpening. The artifacts in the Armadillo’s ears are probably due to the inconsistencies in the loss function of NFGP which asks for \((g_\phi - g_\phi)^2\) and \((\kappa_\phi - \beta \kappa_\phi)^2\), thus the level sets would try to evolve but \((g_\phi - g_\phi)^2\) forces it to be constant.

We also compare with NIE [25]. Given a time step \(\Delta t\), it fits the solution of Eq \((\ref{eq:mean_curvature})\) at times \(t_i = i \Delta t\) using the approximation \(f_{i+1} = f_i - \Delta t \alpha (\nabla f_i, \kappa_i N)\). Thus, for a network \(g_{\phi_i} \approx f_i\), NIE trains the next state \(g_{\phi_{i+1}}\) by minimizing \((g_{\phi_{i+1}} - f_{i+1})^2\). Moreover, a discrete Laplacian estimates \(\kappa_i N\) at the vertices of a mesh approximating \(g_{\phi_i}^{-1}(0)\). The resulting networks \(g_{\phi_i}\) have the domain in \(\mathbb{R}^3\) while we use a single network with a domain in \(\mathbb{R}^3 \times \mathbb{R}\).

Fig 8 shows smoothings of the Armadillo using our method, NIE, and NFGP. Line 1 repeats the results of Fig 5. Line 2 gives the results for NIE using 4/7 steps with \(\Delta t = 1\).

Line 3 presents the results using NFGP to train a network with \(\beta = 0.95, 0.8\) defined empirically. To give a fair comparison, we consider the initial network to be \(f_0(\cdot, 0)\); \(f_0\) is the network of our experiment. We use the procedure in Sect. 3 of the Supp. Material to extract \(f_0(\cdot, 0)\) from \(f_0\).

Importantly, we observe no high-frequency noise in the network derivatives. We used the SIREN extension from [27] that forces the alignment between the surface normals and the network gradient, differently from NFGP and NIE. On the other hand, we observed that computing the mean curvature using the cotangent Laplacian (as in NIE) of a mesh output of a marching cube results in a noisier (see Line 1 of Fig 9) compared to our closed form approach (see Line 2 of Fig 9) using automatic differentiation. The first line of Fig 9 uses the discrete mean curvature (cotangent Laplacian, also used in NIE). Autodiff is applied in the second line without complications.
Comparison with mesh-based approaches

We provide a comparison of our method with a mesh-based approach. We consider the implicit fairing method proposed by Desbrun et al. [15]. Figure 9 shows the results considering the implicit fairing (top) and our approach (bottom) to evolve an Armadillo (with 590k vertices for Desbrun’s method) using the mean curvature flow, which is the parametric version of the mean curvature equation. The colors illustrate the mean curvature. For the discrete case (Line 1) the curvatures are computed using a mesh-based approach: the cotangent formula of the Laplacian of the mesh, also used in NIE. We run the implicit fairing for 100 steps and extracted three Armadillos.

Figure 9. Comparison of our smoothing approach with the implicit fairing [15]. First line shows the implicit fairing smoothing of the Armadillo triangle mesh. Second line illustrates our approach for smoothing a SDF of the Armadillo. Observe that the mesh-based approach can not properly handle regions with high curvature (such as the Armadillo’s fingers). The mean curvature flow should contract those regions like our method does.

Evolving the triangle mesh under the mean curvature flow may lead to shrinkage in high curvature regions where the surface should contract (see the Armadillo fingers). This results in numerical instabilities since the mesh becomes singular in such regions. Another problem is the change of topology during the surface evolution which may create singularities. Mesh-based methods may encounter problems, such as non-manifolds, in regions affected by topology changes (see the blue region in Fig 10). Our implicit approach overcomes this as shown in Fig 4.

Training time for our method is comparable to implicit fairing’s solution time, which took 14.4s per iteration and 24.1 min in total to smooth the Armadillo for 100 steps. Our method took 10.1 min for training, allowing instant evaluation of the smoothed Armadillo without iterative re-runs.

Our approach is also faster than NIE, which required 112 min for 100 time steps (67s per step) for the Armadillo case. In contrast, our method trained on the entire interval in 10.1 min using standard initialization. In Section 6.4 (and in the supp. mat.) we give experiments showing that our initialization scheme allows even faster network convergence.

Finally, observe that our smoothing approach gives a multi-resolution representation of the initial surface. Thus, future works include incorporating multi-resolution neural networks [22, 31] to our framework to fit solutions of the mean curvature equation.

6.3. Interpolation between implicit surfaces

Suppose \( g_i \) are the SDFs of the Bob and Spot (Fig 11, left-right) and that \( f_\theta \) has 1 hidden layer \( f_i : \mathbb{R}^{128} \rightarrow \mathbb{R}^{128} \). We train \( f_\theta \) using \( L_{\text{LSE}} + L_{\text{data}} \), as described in Sec 5.3. Line 1 in Fig 11 shows the reconstructions of the level sets.

Figure 11. Interpolation between Bob and Spot. Line 1 shows the result using our method, and line 2 the Lipschitz MLP. Notice that our method results in smoother transitions between the images.

We compare our method with Lipschitz MLP [23] which considers the tanh activation. We use a Lipschitz MLP with 5 hidden layers of 256 neurons. Each layer is followed by a Lipschitz regularization. The network was trained during 100000 epochs. See the resulting interpolation in Line 2 of Fig 11. Our network is significantly smaller than the Lipschitz MLP, but results in natural interpolation. This is due to the high representation capacity of sinusoidal MLPs.

Figure 12. Interpolation between Witch and Falcon. Line 1 (20000 epochs) is the result of our method and Line 2 (100000 epochs) considers the Lipschitz MLP network. Notice that our approach results in a better approximation of the initial conditions, as can be seen in the Witch’s hood and sword, and Falcon’s beak and talons.
Fig 12 shows the reconstructions of interpolation between the Witch and Falcon (from the Thingi10K dataset [38]). The first line is the result of our method using a network with 2 hidden layers $\mathbb{R}^{128} \rightarrow \mathbb{R}^{128}$ trained during 20000 epochs. For the Lipschitz MLP, we have to consider a larger network with 5 hidden layers of 512 neurons and train it for 100000 epochs. Even with the added capacity and training iterations, the Lipschitz MLP cannot adequately approximate the initial conditions.

6.4. Initialization based on trained networks

Let $g_\theta : \mathbb{R}^3 \rightarrow \mathbb{R}$ be a trained network with 2 hidden layers $g_i : \mathbb{R}^{256} \rightarrow \mathbb{R}^{256}$ that fit the SDF of the Bunny. Let $f_\theta : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}$ be a network with 2 hidden layers $f_i : \mathbb{R}^{256} \rightarrow \mathbb{R}^{256}$. We train $f_\theta$ to approximate a solution of the mean curvature equation subject to $f_\theta = g_\theta$ on $\mathbb{R}^3 \times \{0\}$.

Here we use the scheme of Sec 4.4 to define $\theta$ in terms of $\phi$ such that $f_\theta(p, t) = g_\phi(p)$. We compare it with the standard initialization of sinusoidal MLPs [33].

![Figure 13. Network initialization comparison. Line 1 shows the results of using the initialization is given in [33] and Line 2 considers the proposed initialization. Notice how the new network initialization results in a model that can represent higher frequencies, as shown by the increased surface details in Line 2.](image)

Fig 13 gives qualitative comparisons between the initializations, after training $f_\theta$ during 500 epochs. We empirically observed that our approach speeds up learning. For example, Line 1 shows the bunnies at $t = -0.2, 0.0, 0.6$. Line 2 gives the analogous results using our initialization, which results in faster convergence. Note the preservation of surface details using the proposed initialization at $t = -0.2, 0.0$ (Line 2), compared to the standard initialization in Line 1. Fig 14 shows the plots of the constraints considering 1500 epochs. Note that, using our scheme, the training starts closer to a minimum of the loss function.

![Figure 14. Loss function comparisons. Column 1 shows the plots of the LSE/data constraint using the standard initialization. Column 2 shows the corresponding plots applying our initialization scheme. The horizontal axis represents the number of epochs.](image)

7. Conclusions and Future Work

We introduced a framework to explore the differentiable properties of smooth networks in the problem of evolving level sets of neural SDFs. For this, we extended their domain to space-time and directly operates on primitive neural implicit, which opens up possibilities to control geometric animation and modeling using LSEs, enabling the modeling of multiple surface evolutions in a single method.

The method allows for evolving neural implicit surfaces under LSEs without the use of additional data, only the initial conditions are used. Note that other methods compute an approximation of the solution using numerical simulations and then fit it into the neural network. However, our framework was able to learn the animation only considering the LSE constraint. This is powerful because the models are expressed in compact LSEs that are used to define constraints. This approach enables learning the corresponding animations without any supervision. Most techniques in geometry processing use differential equations to model various kinds of phenomena which, in general, are written in terms of their energy formulation.

The resulting networks are smooth approximations of LSE solutions. Traditional numerical solutions are discrete, making non-trivial the task of introducing more conditions. However, this is a quite simple task in our method. We believe that the development of such methods in graphics would enable the community to use the robustness of classical continuous theories without the need for discretizations.

Acknowledgments

We are grateful to the reviewers for their detailed comments. We thank the Stanford Computer Graphics Laboratory for the Bunny and Armadillo models, and Keenan Crane for the Spot and Bob models. The authors thank CNPQ and FAPERJ for financial support.
References


