Path Routing Optimization for STM Ultrasound Rendering

Hector Barreiro, Stephen Sinclair and Miguel A. Otaduy

Abstract—Ultrasound transducer arrays are capable of producing tactile sensations on the hand, promising hands-free haptic interaction for virtual environments. However, controlling such an array with respect to reproducing a desired perceived interaction remains a challenging problem. In this work we approach this problem as a dynamic mapping of virtual interactions to existing control metaphors of ultrasound devices, namely, the modulation of focal point positions and intensities over time, a method known as Spatiotemporal Modulation (STM). In particular, we propose an optimization approach that takes into account known perceptual parameters and limitations of the STM method. This results in a set of focal point paths optimized to best reconstruct an arbitrary target pressure field.

Index Terms—Ultrasound, haptics, spatiotemporal modulation.

I. INTRODUCTION

THE advent of non-contact haptic displays has introduced new forms of interaction, allowing users to experience tactile sensations in mid-air without the need for holding or wearing a device. Ultrasound haptic devices are notable examples of these displays. They consist of arrays of ultrasound transducers, whose activation is modulated to aggregate pressure waves at specific focal points in space, and thus produce tactile stimulation [1].

The generation of tactile percepts using ultrasound devices is still a largely unknown process. Pressure waves produce a complex mechanical interaction on skin, both in space and time, and this mechanical interaction produces an at least equally complex activation and aggregation of mechanoreceptor signals to form tactile percepts. In the absence of a computational model that maps activation patterns of transducers to tactile percepts, previous works have explored different high-level metaphors to command ultrasound devices. To date, two control metaphors prevail: amplitude modulation (AM) and spatiotemporal modulation (STM). AM controls the position and pressure intensity of focal points to produce pressure distributions on skin, while STM controls paths of focal points to *draw* shapes on skin. In both cases, the high-level control metaphor of focal points is translated into low-level control of transducer activation patterns through well-established optimization methods [2], [3].

We study the problem of rendering interactions with virtual environments using ultrasound haptics. We approach the problem as a dynamic mapping of virtual interactions to the control metaphors of ultrasound devices. Most previous works have simplified this problem by displaying contact locations at maximum intensity, either through AM or STM. On the other hand, we believe that richer display is possible if we account for the force distributions in virtual interactions, and not just contact locations. In our previous work [4] we followed this idea, extracting a target pressure field from the interaction of a virtual hand with a dynamic fluid simulation, and then optimizing AM commands to induce a best-matching pressure field on the user.

However, AM suffers some limitations; probably the most evident that the intensity of focal points modulates a pressure wave that induces a perceivable vibration on skin (typically at 200 Hz). Moreover, STM promises the ability to cover larger skin areas with higher perceived intensity, by leveraging constructive interference of focal

H. Barreiro, S. Sinclair and M. Otaduy are with Universidad Rey Juan Carlos, Madrid, Spain.
Preprint. Author version.

point paths with the skin waves they induce. To date, all previous work commands STM with focal point paths of constant intensity; no previous method computes STM paths of varying intensity to best match dynamic interactions. Not surprisingly, STM poses a more complex challenge than AM. While AM rendering can be regarded as a quasi-static problem, STM requires the solution to a spatial and temporal problem.

In the current work, we propose path routing optimization for STM (PRO-STM), the first method that commands ultrasound STM to render the force distribution resulting from a dynamic virtual interaction. As we discuss in Section III, a key aspect of our method is to pose STM rendering as a quasi-static problem. Thanks to careful approximations, we can eliminate the temporal variable on each dynamic rendering update. As a result, given a target pressure field, we pose STM rendering as the computation of focal point paths that produce the best-matching quasi-static pressure field.

Then, given a target pressure field, we propose an optimization algorithm to compute focal point paths, as described in Section IV. Our algorithm works at two scales. First, on a coarse scale, it initializes paths over the target domain to optimize coverage weighted by pressure intensity. Then, on a finer scale, it refines the paths to maximize the similarity to the target pressure.

We have applied PRO-STM to the interaction with gaseous fluid media, as shown in Fig. 1. In such interactions, haptic perception is dictated by a spatially and temporally varying pressure field on skin, which is used as target for our algorithm. We have compared the reconstruction quality of PRO-STM vs. our previous AM rendering method [4], observing that PRO-STM succeeds to provide larger and smoother coverage than the AM-based method. In particular, AM produces ambiguous results when rendering interaction with one wide plume or with multiple thin plumes, while STM does not suffer such ambiguity. We have conducted an experiment that confirms this observation.

II. RELATED WORK

Iwamoto et al. [1] first demonstrated the capability to use phased array ultrasound to focus sufficient acoustic radiation pressure to induce a haptic sensation in a localized area on the hand. Subsequently, Hoshi et al. [2] extended this achievement to vary the focal location over time, producing the sensation of a moving stimulus. Although the induced skin deformation is slight, haptic detection is ensured by modulating the intensity of the field at a frequency to which the skin's mechanoreceptors, principally the Pacinian corpuscles, are sensitive, approximately 150 to 250 Hz. Considering the focal point location as stable relative to this frequency range, this method later became known as *amplitude modulation*.

These early works used a precomputed time-varying solution for ultrasound phases given a desired spatial amplitude distribution. Because these phases must be determined by means of a non-linear quadratic optimization procedure, the calculations were not amenable to real-time control. Inoue et al. [5] compared various solutions from the field of scattering diffraction imaging, while Long et al. [3] designed an efficient solution that takes advantage of the linearity of the complex-valued eigenproblem associated with the desired focal point intensities.
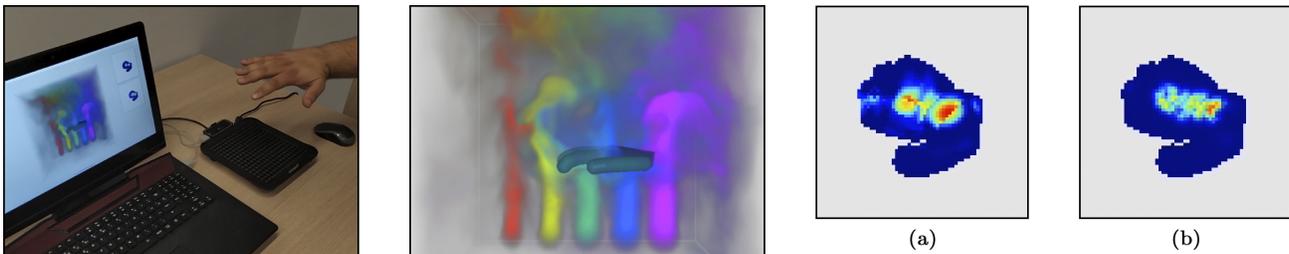
Fig. 1. Example scene rendered using our novel PRO-STM method. From left to right: Physical setup with the ultrasound device and a view of an interactive fluid simulation; screen-capture of the fluid simulation, showing the virtual representation of the user's hand interacting with colored smoke plumes; pressure field (a) produced by the smoke on the surface of the hand, which sets the target for our algorithm; reconstructed pressure field (b) produced by STM rendering of dynamically optimized focal point paths.

Later, Korres and Eid [6] pointed out that the efficiency of this solution admits a different display method, in which focal point amplitudes are relatively stable but their location is modulated at much higher rates. This was dubbed *spatiotemporal modulation* by Kappus and Long [7], and they showed that such paths could produce recognizable tactile shapes.

Since, as opposed to AM, the STM display cannot be thought of as coupling to specific frequencies, new investigations were needed to understand the tactile perception induced by this method. Therefore, recently more studies have been performed investigating various aspects of STM from perceptual and biomechanical points of view. Frier et al. [8] found that there is some optimal focal point speed (between 5 and 8 m/s in their experiments) to maximize skin sensitivity. Recently, they have analyzed the combined influence of spatial and temporal sampling in STM [9]. Takahashi et al. [10] proposed a method that falls in between AM and STM, wherein they use a spatial modulation of the constant-intensity focal point at 1000 Hz to produce a vibration signal in a localized range. They report lower detection thresholds than for a single AM point. Howard et al. [11] investigated the detection and discrimination of line patterns with respect to intensities, as well as the discrimination of bumps and holes based on varying intensity over a line. They also found that STM line patterns had a lower detection threshold than a single AM point. Finally, Reardon et al. [12] captured the wave propagation induced in the palm by ultrasound haptics using an optical vibrometer, and found that waves above 4 m/s induced compression. In this regime, it was found that inhibiting these waves reduced motion perception, thus a connection between wave propagation in the skin and motion perception was inferred.

In our demonstrations, we apply PRO-STM to the interaction with a gaseous medium because it provides a good example of dynamic pressure field, and the tactile stimulus is not too intense; hence it can be reasonably matched even under the power limitations of current ultrasound devices. There are methods designed specifically for rendering air flows [13], but our algorithm supports arbitrary interactions and makes no assumption on the source of air flows.

III. PRINCIPLES OF SPATIOTEMPORAL MODULATION

To design an ultrasound rendering algorithm based on STM, it is important to understand the properties that best balance the capabilities of the ultrasound device with the quality and richness of tactile stimuli. We pay attention to the speed and frequency at which focal points traverse skin, and we extract desiderata for our algorithm. In addition, we devise a model of the radiation pressure produced by focal point paths, which helps design our algorithm. Specific values of the parameters and constraints of the algorithm depend on the choice of ultrasound device, an Ultrahaptics STRATOS device in our case.

A. Rendering Parameters and Constraints

In STM, a focal point traverses a path in space. We formally describe the path as a time-dependent position: $\mathcal{P} = \mathbf{x}(t) \in \mathbb{R}^3$. As noted earlier, Frier et al. [8] concluded that the perceived intensity of a focal point path is maximized under constructive interference between the motion of the focal point and the propagation of skin waves. This happens for focal point speeds between 5 and 8 m/s; therefore, we select a reference speed $v = 7$ m/s for our rendering algorithm. Furthermore, to ensure constructive interference on the complete path, we design closed paths, i.e., $\mathcal{P}$ is a closed 3D curve.

Frier et al. assumed that the length of the path is given; therefore, the frequency at which the path is repeated cannot be independently controlled, and depends on the traversal speed and the length of the path. In our preliminary experiments, we observed that this is acceptable up to a maximum path length. Beyond that length, the frequency at which the path is repeated is too low, and the stimulus is no longer perceived as continuous. To determine the minimum acceptable frequency, i.e., the maximum acceptable path length, we informally experimented rendering circles of different radii at the reference traversal speed of 7 m/s. We found that a minimum frequency of 50 Hz, i.e., a maximum path length $L = 140$ mm is a safe bound to ensure that the stimulus is perceived as continuous.

Our test device allows STM of multiple focal points simultaneously. Each focal point can traverse a different path, with all focal points traveling at the reference speed, and all paths satisfying the maximum length constraint. From our preliminary experiments, we have concluded to limit the number of simultaneous focal points, i.e., the number of simultaneous paths, to four. More focal points may reach larger coverage, but at the price of notable degradation of perceived intensity.

B. Quasi-Static Pressure Field

In previous work, STM was used to render 3D curves, hence the intensity of the radiation pressure of the moving focal point was kept constant along such curves. In our work, we render a temporally and spatially varying pressure field, hence the intensity of the radiation pressure along the path should adapt locally to the intensity of the pressure field. Based on this consideration, we characterize a path with a position-dependent pressure intensity $p(\mathbf{x})$.

As a focal point cycles multiple times through the same position $\mathbf{x}_i$, the rendered pressure intensity $p(\mathbf{x}_i) = p_i$ is the same on all cycles. If the path is repeated frequently enough (i.e., at more than 50 Hz), the rendered radiation pressure produces a persistent tactile perception. We consider that this is equivalent to applying a time-invariant pressure at each position along the path, with its effective magnitude a fraction of the rendered pressure. Thanks to this assumption, during a time window we can consider that the effective

pressure is a spatially varying but temporally invariant field, i.e., a *quasi-static* pressure field.

Determining accurately the effective pressure of STM rendering is a complex subject that requires further research. It is not a simple time-average of the rendered pressure. In our work, we follow a perceptual heuristic to approximate its magnitude. We render the same stimuli using AM and STM, and we ask subjects to tune the gain γ of STM until the peak perceptual intensity is similar. In practice, we have used a gain $\gamma = 1.4$.

Focal points exhibit a smooth fall-off determined by the wavelength of the ultrasound signal (e.g., 8.6 mm for the 40 kHz of our test device). As shown by Hoshi et al. [2], this fall-off can be approximated well by a Gaussian function ϕ. Based on this finding, together with the heuristic gain γ, we approximate the effective quasi-static pressure field produced by a focal point path as

$$p(\mathbf{x}) = \frac{1}{\gamma}\, p_i\, \phi(\|\mathbf{x} - \mathbf{x}_i\|) = \frac{1}{\gamma}\, p_i\, e^{-\frac{\|\mathbf{x} - \mathbf{x}_i\|^2}{2\,\sigma^2}}, \qquad (1)$$

where $\mathbf{x}_i$ is the closest position to $\mathbf{x}$ in the path. We set the standard deviation σ of the Gaussian fall-off to the same value as the wavelength of the ultrasound signal (i.e., 8.6 mm)

The quasi-static pressure field assumption simplifies the design of our rendering algorithm. Given a target pressure field obtained from a fluid simulation, we pose each rendering step as the search for the focal point paths whose quasi-static pressure field best reconstructs the target field. This search must fulfill two constraints to ensure that the quasi-static pressure field assumption is valid, namely that each focal point travels at the reference speed v and the length of each path is not longer than the maximum length L. In the next section, we describe our path optimization algorithm.

IV. PATH ROUTING OPTIMIZATION

Given a target pressure field, we seek focal point paths that produce a best-matching quasi-static pressure field. We solve this problem in two steps, at two different resolutions. Both steps search for paths that maximize coverage and integrated pressure intensity subject to the path length constraint, but the first coarse step performs a global search, while the second fine step performs a local search. Before describing these two steps in detail, we describe how we obtain the target pressure field from a fluid simulation. And we conclude the section with implementation details to render the paths on our test device.

A. Target Pressure Field

We adopt the approach in [4] to compute an interactive fluid simulation and extract the target pressure field. Same as that work, we track the user's hand and model it as a moving obstacle in a 3D simulation of a gaseous medium. We model fluid dynamics using incompressible Euler equations discretized on a 3D Eulerian grid, with semi-Lagrangian advection and massively parallel Jacobi relaxation for the pressure solve. The fluid simulation is executed on a GPU for maximum performance. Please refer to [4] for full details.

To define the target pressure field, we voxelize the hand, and select the voxel positions that are visible from the side of the domain that corresponds to the location of the ultrasound device. To simplify the path optimization problem, we fit a plane to the voxel positions and we project them onto the plane, making path optimization a 2D problem. Formally, the target pressure field is described by a set of 2D positions and their corresponding target pressure values, $\mathcal{T} = \{(\mathbf{x}_i \in \mathbb{R}^2, p^*(\mathbf{x}_i))\}$.

To ensure high computational performance, the initialization of the paths uses only a representative subset of the target pressure points $\mathcal{T}$. We apply weighted k-means clustering to $\mathcal{T}$, to produce a downsampled target pressure field with pressure points $\mathcal{D}$. In Fig. 2-a we show $\mathcal{T}$, the target pressure points on the user's hand extracted from a fluid simulation, and in Fig. 2-b we show $\mathcal{D}$, the downsampled pressure points.

B. Path Initialization

Given the pressure points $\mathcal{D}$, we seek a set of closed paths that visit all the points, subject to the maximum path length L. The optimal solution to this problem may require an arbitrarily large number of paths; however, as noted in Section III-A, we concluded to limit the number of paths to four in practice. Consequently, the resulting paths may fail to visit all the pressure sample points, and an optimal set must be selected.

We solve this problem iteratively. We first compute the optimal path that visits all the points. If the path is too long, we split the set of points into two subsets and we compute separate optimal paths. We split the subsets of points recursively until all paths satisfy the maximum length constraint. Fig. 2-c shows the optimal path for the full set of pressure sample points, while Fig. 2-d shows the optimal paths after splitting the points to satisfy the maximum path length constraint. If the number of resulting paths is larger than four (as in the figure), we retain the four paths with highest integrated pressure, and we pass them to the refinement step described in the next section. But first, we describe in detail the operations to compute an optimal path for a set of points and to split a set of points.

Given a set of points, finding the shortest path that visits all the points corresponds to the *traveling salesman* problem. We solve this problem using the 2-opt algorithm [14], which admits closed paths. The computational cost of 2-opt sets an upper bound on the size of $\mathcal{D}$. In our implementation, we set it to a maximum of 50 points. Thus we run the weighted k-means clustering step above with 50 or fewer clusters.

To split a set of points, we find the direction of maximum spread, we bound the points along this direction, and we place a splitting plane orthogonal to the direction at the midpoint of the two bounds. To find the direction of maximum spread, we compute the covariance matrix of the points, weighted by their pressure value. The direction of maximum spread corresponds to the eigenvector with highest eigenvalue.

C. Path Refinement

After initialization, the paths pass through pressure clusters and fulfill the maximum length constraint. However, due to their coarse sampling, they are not optimally aligned with pressure peaks and ridges. We execute path refinement at a higher resolution, hence we start by upsampling each path to N points. $N = 20$ in our implementation, which sets points 7 mm apart from each other, i.e., the distance traveled by a focal point in 1 ms.

During refinement, the goal is to move path samples locally toward locations with higher pressure, while ensuring that paths preserve the following properties: (i) they satisfy the maximum length constraint; (ii) to reach maximum coverage, they do not (self-)intersect; and (iii) they do not bend at sharp angles, as the design decisions of our algorithm stem from perceptual observations on smooth paths, and paths with sharp corners reach smaller coverage. To implement refinement, we formulate the goal and the properties as cost terms of an objective function, and we execute a minimization algorithm.

Given paths with samples $\{\mathbf{x}_i \in \mathbb{R}^2\}$, we formulate a pressure intensity cost term as:

$$c_p = -\sum_i p^*(\mathbf{x}_i). \qquad (2)$$
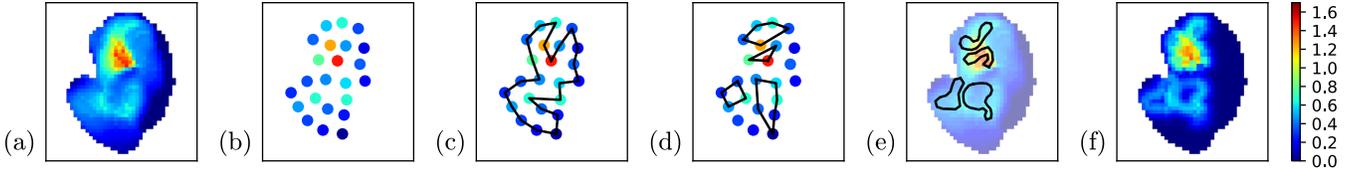
Fig. 2. Steps of our PRO-STM algorithm: (a) input target pressure field, (b) clustering, (c) initial path, (d) split into multiple paths to satisfy length constraints, (e) refinement of the paths to maximize pressure intensity, (f) resulting reconstructed pressure field.

This term is minimized as the samples move to locations with higher pressure.

With a target path length L, and N samples per path, the target length is obtained if the length of each path segment is L/N. Then, we formulate a length cost term as:

$$c_l = \sum_i (\|\mathbf{x}_{i+1} - \mathbf{x}_i\| - L/N)^2, \tag{3}$$

where $\mathbf{x}_i$ and $\mathbf{x}_{i+1}$ are two consecutive path samples.

If two paths or two portions of a path get closer than the fall-off distance of focal points, σ, they stimulate overlapping skin areas. The result can be considered inefficient, as the covered skin area is larger if the paths move away. We introduce a (self-)intersection cost term that prevents path samples from getting too close:

$$c_i = \sum_{i,j} \max(\sigma - \|\mathbf{x}_j - \mathbf{x}_i\|, 0)^2, \tag{4}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are two non-consecutive path samples.

Finally, to favor low-curvature paths, we introduce a bending cost term:

$$c_b = \sum_i \arctan^2 \frac{(\mathbf{x}_{i+1} - \mathbf{x}_i) \times (\mathbf{x}_i - \mathbf{x}_{i-1})}{(\mathbf{x}_{i+1} - \mathbf{x}_i)^T (\mathbf{x}_i - \mathbf{x}_{i-1})}, \tag{5}$$

where $\mathbf{x}_{i-1}$, $\mathbf{x}_i$ and $\mathbf{x}_{i+1}$ are three consecutive path samples.

We optimize the paths by iterating steps of gradient descent of the four cost terms. For the pressure intensity term, we set a 2D grid with the target pressure values $\mathcal{T}$, and we use bicubic interpolation to evaluate the pressure at subgrid resolution and to compute robust gradients. Furthermore, we apply a line search to ensure that the step along the gradient reduces the cost. Fig. 2-e shows example paths after refinement.

To account for the effective magnitude of the pressure field, we must incorporate the heuristic gain γ discussed in Section III-B. We set the rendered pressure of a point on a path as $p_i = \gamma\, p^*(\mathbf{x}_i)$, based on the target pressure field p^*. Fig. 2-f shows the reconstructed pressure field according to the quasi-static pressure model of Section III-B.

D. STM Rendering

Once 2D paths are fully computed, we lift them back to 3D for rendering on the device. We achieve this by undoing the projection of the hand voxels discussed in Section IV-A.

Each path is 140 mm long and is traversed in 20 ms at 7 m/s. The STM rendering API of the Ultrahaptics STRATOS Explore (USX) device updates every 1 ms a burst of 40 consecutive focal point positions. Therefore, we linearly upsample each path to 800 points spaced 0.175 mm, and feed them in groups of 40 points to the API.

V. EXPERIMENTS AND RESULTS

Here we describe some details of implementation and characteristics of its run-time performance. Following this, we describe a user study we performed to compare PRO-STM to the AM method [4].

	1-plume		4-plumes	
	Mean (μs)	StdDev	Mean (μs)	StdDev
Init:	285	106	142	59
K-means:	1548	386	948	338
Initial path (2-Opt):	2550	646	1224	554
Splitting (recur. 2-Opt):	744	235	324	179
Resampling:	12	4	10	4
Refinement:	17896	1664	11120	3344
Total:	**24303**	**2256**	**14640**	**4281**

TABLE I
TIMING FOR EACH STEP OF PRO-STM ON TWO SCENARIOS.

A. Implementation Details

All experiments were performed, and timing data was collected, on a Lenovo laptop featuring an NVidia GeForce GTX 1070 GPU and Intel Core i7-6820HK CPU at 2.7 GHz, with 24 GB of RAM and 8 GB of video RAM. Software was written in C++ and CUDA for GPU kernels.

We have used an Ultrahaptics STRATOS Explore (USX) device for ultrasound rendering. The USX employs an array of 16×16 transducers running at 40 KHz, thus producing focal points of 8.6mm diameter. The device supports a maximum of 8 simultaneous focal points, but we have used 4 in our experiments. Hand pose tracking was achieved through the Leap Motion tracker device bundled with the USX.

The fluid simulation ran at 50 Hz on the GPU while PRO-STM was executed in parallel on the CPU at 25 Hz. At this rendering rate, each path was traversed twice per update.

B. Performance

We evaluated computational performance of PRO-STM on 500 frames recorded during interaction with two scenarios, namely 1 and 4 plumes of smoke, as these have qualitatively different distributions of pressure fields. Timings were broken down by segment of the algorithm, and are presented in Table I. These data show that the algorithm runs at interactive rates (typically above 40 Hz in this implementation) without problems on commodity hardware, and may do so alongside a computationally heavy load driving a time-varying target field. In practice we run PRO-STM at 25 Hz in order to attain integer division of the path traversal frequency of 50 Hz, as discussed above.

We have also compared performance with the AM rendering method [4]. AM shows superior performance, with update rates of roughly 250 Hz on both test scenarios.

C. Qualitative Comparisons

Currently we restrict the comparison between the current work and the previous AM method [4] to a qualitative one because the respective reconstructed fields are not numerically comparable. This is because the AM field reconstruction is very sparse compared to the STM field reconstruction, but it is not clear from current
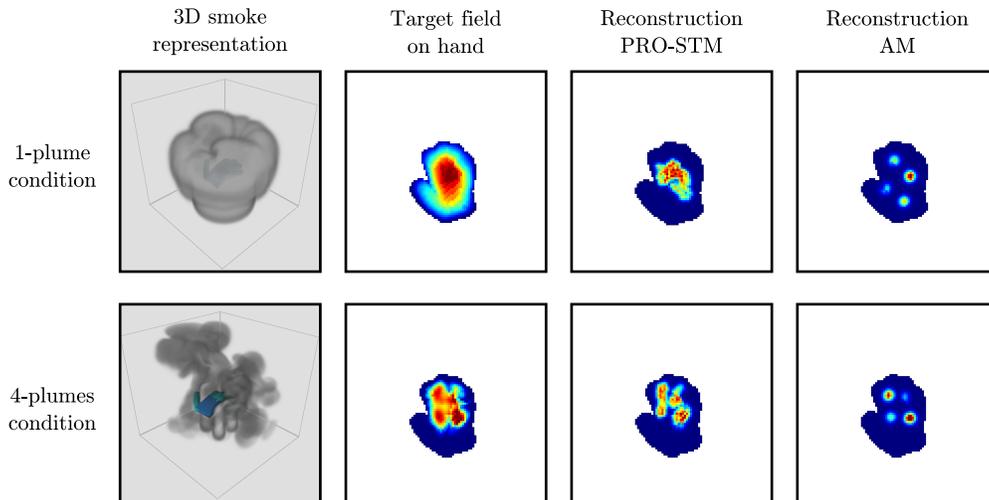
Fig. 3. Experimental scenarios and representative examples of target and reconstructed fields using the proposed method (PRO-STM) and a previous AM method [4].

knowledge of these haptic display methods how to properly consider the reconstructed fields from a perceptual point of view. In particular, the relation between the rendered pressure intensity and the perceived intensity is different for these two methods, and this variable strongly affects computed reconstruction error. The reconstruction model does not describe well the effects of parasitic tactile artifacts either, such as AM's constant vibration.

However, to demonstrate qualitatively the difference in coverage and detail reproduction, in Fig. 3 we visualize reconstructions of rendered pressure fields with STM and AM. For STM, we use the model of quasi-static pressure field of Eq. 1, which includes the heuristic gain between rendered and effective pressure. Both for STM and AM, we use $\sigma = 8.6$ mm, corresponding to the focal point fall-off. Since our model of quasi-static pressure reconstruction is based on heuristics, it is not suited for a quantitative (e.g., RMSE) comparison of STM and AM.

The renderings are taken from the scenarios used in the comparison study described later in this section. It is apparent that the field expected from STM does a better job at differentiating between these two conditions, while AM suffers from ambiguity, making it difficult to tell the difference between one large smooth shape and four smaller ones. In the following section, we describe a study intended to investigate this ambiguity and whether PRO-STM might overcome it.

D. Perceptual Study

To evaluate a difference in rendering between the proposed algorithm and previous AM work for fluid rendering [4], we explored whether PRO-STM could provide more details to the user in a situation where the AM approach leads to ambiguity in the display.

Specifically, we noticed that when a large, smooth pressure field is present, such as in the case of a single, large plume of smoke rising towards the hand, the AM approach places four control points in relatively stable positions. Correspondingly, it can be difficult to discriminate this display from a situation of four individual pressure concentrations, as in the case of four smaller columns of smoke. These situations are depicted in Fig. 3.

In the study, we asked participants to discriminate between 1-plume and 4-plume scenarios. The study featured 8 participants
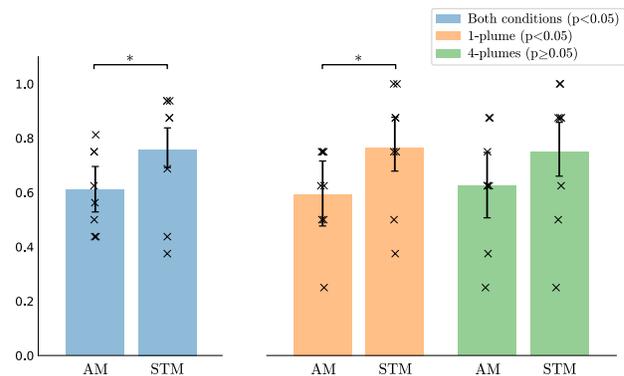
Fig. 4. Proportions of correct responses for the two methods overall, and by scenario. Asterisks represent significant differences, determined by a pooled probability z-test across the cumulative samples. Error bars indicate standard error of participant means, with individual proportions marked by an 'x'.

evaluated with AM, and 8 separate participants evaluated with PRO-STM. Initial testing revealed some confounding effects of mixing the two methods in the same experiment, as participants would use detectable differences between the two methods as erroneous cues to attempt to distinguish the two scenarios. Therefore, we found it most reliable to examine average performance across participant groups each evaluating different methods. No participants from earlier testing participated in the final study.

Participants were seated in front of the STRATOS device and a laptop running the fluid simulation. They were asked to hold their hand out flat above the device and feel the column or columns of smoke. First, an initial training period of two minutes was given where they were allowed to explore freely the two scenarios while seeing the simulation's 3D visual representation. Next, each participant performed 16 trials, 8 with one column and 8 with four columns, in random order. On each trial, the participant experienced the scenario while looking at a static image along with a textual question, chosen randomly, "Is there one plume of smoke?" or "Are there four plumes of smoke?" The experimenter also repeated this text out loud for each trial, and was responsible for entering the information. The participant responded with "true" or "false" by

voice. Participants did not have an explicit time limit to experience the scenarios and respond, but they took on average 5 seconds per trial.

Demographics: participants were between the ages of 20 and 50, with two participants above 40. Seven participants were female, and 9 were male. The majority of participants interacted with the experiment in their native language, Spanish. Two participants performed the experiment in English. One participant was left-handed and performed the experiment with his dominant hand. No participants reported health difficulties, abnormal skin conditions, or sensitivity issues.

A significant difference was found between AM and STM, with higher proportion of correct responses using STM, see Fig. 4, indicating that PRO-STM delivers more detailed information about the target in both scenarios. No effect of the scenario was found, and since participants tried only one method each, no order of presentation effect is possible. Additionally no significant difference was found between genders or due to question correctness. The same experimenter performed all trials, which took place in an office environment.

Notable in the figure is that some participants could not perform the task, whereas others performed it exceptionally well. This would suggest a possible bimodal distribution, and we could exclude poorly performing individuals, but we elected not to as it did not affect the significance of the cumulative result. However, we noted that during informal testing and in demo conditions some people had difficulty detecting or discriminating using either AM or STM or both; this appears to be a limitation of ultrasound haptics for which we do not have enough data to correlate with explanatory variables such as skin type, hydration, or sensitivity, and remains an open question for the field.

VI. CONCLUSION

We have described a new method, PRO-STM, for dynamically determining the paths of STM ultrasound focal points for an arbitrary target pressure field, and demonstrated improved coverage and smoothness over previous work based on AM ultrasound.

We account for current knowledge of STM to design algorithm features such as optimal traversal speed, maximum path length, or minimum inter-path distance. However, several perceptual factors related to STM remain unknown, and they can spur further algorithm improvements.

(i) We currently treat path routing as a 2D problem, whose solution is subsequently un-projected to a 3D path, and we ignore the change of path length due to this un-projection. The problem admits a more accurate formulation, accounting for the surface of the hand as a 2D manifold embedded in 3D; however, the effect of focal point velocity in 3D is an unknown aspect of STM perception. (ii) We introduce a bending cost term in the optimization, as we know that path corners are difficult to distinguish; however, we do not know how to best weigh this cost, as the sensitivity of STM perception to changes in direction is unknown. (iii) We optimize paths for each rendering update independently, without special treatment to path transitions. This approach induces transients with unknown effects, as moving focal point paths have not been studied yet. (iv) We map the rendered pressure to the effective perceived pressure using a heuristic gain. We use a constant gain, but in reality the mapping depends on the spatiotemporal characteristics of each path and the specific effect produced on the subject's skin; effective skin deformation might be a more accurate metric to design a mapping. (v) The implementation of the algorithm relies on the choices of optimal traversal speed of 7 m/s and minimal frequency of 50 Hz. While these values are derived from perceptual studies, they can be refined through further experiments focused on the algorithm.

Although recent works are beginning to address perception of STM and its relation to wave propagation in the skin [12], we lack knowledge of the effect on perceived intensity of fundamental STM parameters such as rendered intensity, traversal speed, or path frequency. This knowledge can improve our quasi-static pressure reconstruction model, and allow a more considered path optimization. Overall, as our method depends on hand tracking, it is in a good position to take into account new knowledge about hand skin mechanics and its connection to ultrasound-based stimulation.

REFERENCES

[1] T. Iwamoto, M. Tatezono, and H. Shinoda, "Non-contact method for producing tactile sensation using airborne ultrasound," in *Proceedings of the 6th International Conference on Haptics: Perception, Devices and Scenarios*, ser. EuroHaptics '08. Springer-Verlag, 2008, pp. 504–513.

[2] T. Hoshi, M. Takahashi, T. Iwamoto, and H. Shinoda, "Noncontact tactile display based on radiation pressure of airborne ultrasound," *IEEE Transactions on Haptics*, vol. 3, no. 3, pp. 155–165, 2010.

[3] B. Long, S. A. Seah, T. Carter, and S. Subramanian, "Rendering volumetric haptic shapes in mid-air using ultrasound," *ACM Transactions on Graphics (TOG)*, vol. 33, no. 6, p. 181, 2014.

[4] H. Barreiro, S. Sinclair, and M. A. Otaduy, "Ultrasound rendering of tactile interaction with fluids," in *2019 IEEE World Haptics Conference (WHC)*, 2019, pp. 521–526.

[5] S. Inoue, Y. Makino, and H. Shinoda, "Active touch perception produced by airborne ultrasonic haptic hologram," in *2015 IEEE World Haptics Conference (WHC)*, 2015, pp. 362–367.

[6] G. Korres and M. Eid, "Haptogram: Ultrasonic point-cloud tactile stimulation," *IEEE Access*, vol. 4, pp. 7758–7769, 2016.

[7] B. Kappus and B. Long, "Spatiotemporal modulation for mid-air haptic feedback from an ultrasonic phased array," *The Journal of the Acoustical Society of America*, vol. 143, no. 3, pp. 1836–1836, 2018.

[8] W. Frier, D. Ablart, J. Chilles, B. Long, M. Giordano, M. Obrist, and S. Subramanian, "Using spatiotemporal modulation to draw tactile patterns in mid-air," in *Haptics: Science, Technology, and Applications*, D. Prattichizzo, H. Shinoda, H. Z. Tan, E. Ruffaldi, and A. Frisoli, Eds. Springer International Publishing, 2018, pp. 270–281.

[9] W. Frier, D. Pittera, D. Ablart, M. Obrist, and S. Subramanian, "Sampling strategy for ultrasonic mid-air haptics," in *CHI Conference on Human Factors in Computing Systems Proceedings*. ACM, 2019.

[10] R. Takahashi, K. Hasegawa, and H. Shinoda, "Lateral modulation of midair ultrasound focus for intensified vibrotactile stimuli," in *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*. Springer, 2018, pp. 276–288.

[11] T. Howard, G. Gallagher, A. Lécuyer, C. Pacchierotti, and M. Marchal, "Investigating the recognition of local shapes using mid-air ultrasound haptics," 2019.

[12] G. Reardon, Y. Shao, B. Dandu, W. Frier, B. Long, O. Georgiou, and Y. Visell, "Cutaneous wave propagation shapes tactile motion: Evidence from air-coupled ultrasound," in *IEEE World Haptics Conference (WHC)*, Tokyo, Japan, July 2019, pp. 628–633.

[13] R. Sodhi, I. Poupyrev, M. Glisson, and A. Israr, "Aireal: Interactive tactile experiences in free air," *ACM Trans. Graph.*, vol. 32, no. 4, pp. 134:1–134:10, 2013.

[14] G. A. Croes, "A method for solving traveling-salesman problems," *Operations Research*, vol. 6, no. 6, pp. 791–812, 1958.