# Proceedings of CSCI2370'20

## **Interdisciplinary Scientific Visualization**

Prof.: David H. Laidlaw TA: Fumeng Yang

**1. Towards Semantically Meaningful Visualization Tools for Choreography and Improvisation in Dance** 

Megan Gessner

**2. Multidimensional Animated Visualizations for Transgender Voice Training** Luna McNulty

# **3.** Visualizing Alzheimer's Disease Brain Pathology Using Deep Conditional Population Templates

Shuyan Wang

**4. Visualization Methods to Compare Metrics to Quantify Forest Structure** Christopher Wong

> December 2020 Brown University

## Towards Semantically Meaningful Visualization Tools for Choreography and Improvisation in Dance

Megan Gessner\*

**Brown University** 

#### ABSTRACT

This report shares insights and open problems that I have found from exploring the literature and various existing systems produced by the movement computation community which analyze, visualize, generate, or annotate motion. The state of the art is moving towards real-time, fully-immersive choreographic interfaces that can intuitively represent the high-dimensional feature space of movement and allow us to infer physiological experiences, relationality to technique, experience, and community, and expression of conscious and subconscious cognitive states through movement - an expansion of knowledge that would push the boundaries in both art and science. Selected from a larger set of design principles, motivations, technologies, and methodologies that have shaped this interdisciplinary field, this report highlights those findings from the research that may be relevant for grounding future work in developing novel interactive embodied visualization tools that support dance education as well as the creative practice of improvisers and choreographers. Finally, I posit that the identified open areas will necessitate an integration of state-of-the-art solutions across scientific visualization, computer vision, AI, biomechanics, pattern recognition, cognitive science, and dance studies, positioning this inquiry as an interdisciplinary science.

#### **1** INTRODUCTION

Dance and computation have been working as partners in research for a couple decades now. Their partnership has motivated the notion that movement can be formalized as a multivariate, continuous function on the body over space and time. In general, dance is concerned with the expressivity of movement, and how our corporeal existence reflects our individual and collective experiences. Choreographers and dancers explore the multitude of ways in which movement can vary: effort, shape, trajectory, rhythm, musicality, emotion, relationality, speed, stillness, energy, posture, gesture, narrative, the list goes on. One of the aims of exploratory improvisation and choreography in dance is to seek out new movement ideas and bodily experiences, and find clarity in the communication/expression of the introspective self. Interactive visualization tools and computational methods can augment this work of creative movement practitioners towards this end: by presenting analytical insights, providing new ways of improving their understanding of their craft and themselves. Introducing novel representations of movement data may inspire dancers to experiment with how to induce new patterns and visual effects as a function of their movement choices, a form of creative prompting that happens in conversation with the dancer. Ultimately, the significance of the work at the intersection of dance and technology is its power as a research and educational tool for the dance community. [5]

In order to ultimately create a novel visualization from the fusion of dance and computer science, one needs to do two things:

1) understand the body of existing work for context (so as to learn from it, build on it, and not duplicate the results) and 2) identify what problems need solving and could benefit from computational methods and visualization. Regarding the latter, I motivated my research this semester with a desire to bring awareness to an improvising dancer about their movement habits and get them to examine what choices they were making and how similar or unexpected their choices are in relation to past movements. With no prior reading, I imagined a possible solution to this could be a real-time visualization consisting of silhouettes overlayed on a virtual avatar that represented the system's predictions about future movements given current movements. Using this idea as a benchmark, I went on to do a literature search, so I could get some context about whether this has been done and how I might approach the design, and so forth. In doing so, I discovered that a whole community dedicated to "movement computation" is tackling this and a plethora of other problems concerning how technology and computation can be used to understand, augment, analyze, annotate, visualize, or generate movement and provide creative and educational tools for dancers, choreographers, and movement practitioners. Therefore, this report captures an overview of the state of the art of this field, its existing work, and its open questions, to better ground myself and others like me in the knowledge of what is already out there, how its challenges were overcome, and where we can go next using this knowledge.

#### 2 THREE OPEN AREAS FOR DEVELOPMENT

I will discuss three open areas in movement computation and visualization: 1) building tools that are sufficiently flexible and generalized to support all styles and choreographic practices (in contrast to existing narrow-scope applications), 2) understanding and representing movement sequence patterns and their underlying cognitive motivations, 3) artistic movement augmentation with wholebody interaction, specifically using XR to allow for instantaneous response to data and multiple simultaneous views/features. In addition, I will give examples of systems that begin to solve these problems and address their benefits and shortcomings. For a detailed survey of the state of the art in interactive educational and creative tools for dance, I refer the reader to [20, 1, 16].

#### 2.1 Broadening the Scope of Creative Tools for Dance

In my survey of existing tools, I found that many of them worked to provide solutions for a single specific genre, like ballet [14], contemporary, or Greek folk dance, or for a single feature of movement, like evaluating based on teacher-student posture correspondence, or for analysis of a single choreographer and their work, as in [9]. This causes a proliferation of hyper-specific tools meant for a small subset of the dance community, which limits the available movement possibilities for each tool. Furthermore, many tools were only built to be used in an experimental context, without a public software release, or rely expensive motion capture equipment, limiting the accessibility of these tools to the larger community. An important next step is to allow these tools to be accessible and applicable to all, like that of [2], which means building systems that support multiple styles of dance, ability levels, learning approaches, and

<sup>\*</sup>e-mail: megan\_gessner@brown.edu

creative compositional or improvisatory techniques. By centralizing the knowledge and tools generated by movement researchers, we can design modular systems that can adapt to a user's needs, for example, generalizing evaluative frameworks to respond to any genre of input movement phrase, creating a repository of task-based visualizations for prompting movement (similar to that of [19]), or aggregating motion data and annotation interfaces like [20].

#### 2.2 Understanding Movement Patterns

There is room to dive even deeper into the minds of dancers, into the full complexity of the movement without oversimplification, and to analyze movement sequences as a whole over time, rather than with instantaneous measures of qualities. I've observed that many systems contain the same types of visualizations: trails for bone trajectories, colors for effort qualities, body pose as an skinned mesh avatar or stick figure. Although these have been shown to be useful, the complexity of the data features being visualized are a mere fraction of what is possible, and are largely a reflection of the raw pose data itself. Because movement has so many degrees of freedom and is such a large and rich source of data, much work has been done to compress that data [4], efficiently store and retrieve it from motion databases [13, 21], and design similarity measures or clustering techniques based on characteristic features for comparing movements [8, 17]. With these innovations, we can begin to do advanced reasoning about the hidden information within movement: how a movement sequence relates to itself, to others, and to a body of previous dance training/technique, and what this may say about the cognitive processes mediating the choices that are made in improvisation.

One very recent example [10] begins to scratch the surface of this particular domain, where they used a set of five movementcontrolled 3D visualizations to identify a set of six embodied interaction patterns that emerged when dancers improvised with their system. They observed the impression/expression feedback loop that pervades improvisation of sensing the self and the environment and then responding in some meaningful way, and categorized the abstract relationships that are made between dancer and interface. With this knowledge, we could learn to computationally recognize these interaction patterns and gently bring the dancer's awareness to which type of pattern they are engaged in by responsively adjusting or adding to the visualization. I imagine that we could use similar visualization-dancer interaction experiments to study and then apply knowledge about how the perception and cognition affects embodied expression, whether that is through deep-learned recognition of facial affect, gestural interpretation to construct protonarratives, or studying entire movement sequences like genomic strings in order to ascertain what cognitive process brought them to life (habitual, spontaneous, reactive, etc.). With developments like these, we can start to visualize the internal states of dancers in tandem with their movement, which would allow dancers to hone in on a sense of agency and clarity in their movement choices, as well as improve the understanding of the audience/general observer (towards artistic appreciation and communication of ideas).

#### 2.3 Immersive, Reflective, and Analytical Visualization

Visualizations can impact dance making and learning in several ways: the sketching and modeling of new ideas [6], summarizing the structure and composition of a piece of choreography [7], serving as an augmented/abstracted mirror [18] or virtual dance partner [11, 15], cueing, prompting, and generating new movement ideas [12], and training/teaching new dancers in established techniques [3] via virtualized versions of traditional learning methods [16]. As I've already discussed, the existing systems still have quite a ways to go to intuitively represent the complexity of movement data, but the real gap in the work is an effective, efficient interaction scheme that doesn't interrupt the creative process. Visualizing dance on

desktop computer requires either doing all the creative work offline [6] (i.e. not in a studio) or moving back and forth from dance floor to computer to make adjustments, which is not conducive to the state of flow many choreographers seek. I think the most promising avenue for getting to a seamless user experience are immersive systems, i.e. using AR/VR, that provide real time feedback and streams of data about the dancing and make use of 360 degrees of space to display information that the dancer can interact with. And since dance is so inherently communicative, these systems should employ gestural interfaces where a user can move their body to interact with what is visualized or the choices of data visualization respond in real time to choices that the dancer makes.

#### **3** CONCLUSION

The field of movement computation and its intersection with scientific visualization is surprisingly prolific. The research done to design systems that assist choreographers and improvisers in their creative processes and teach and train students of dance raise a lot of questions about the salient features of movement and how collected data should be visualized to best inform artists. As more work is produced, we learn more about how the mind works, how the body extends and transmits mental states, and how to process and represent such high dimensional data, which can impact other fields with similarly complex data.

#### **A**CKNOWLEDGEMENTS

The author wishes to thank her professor David Laidlaw, her TA Fumeng Yang, and her classmates Shuyan Wang, Luna McNulty, and Chris Wong, for making this class an unforgettable learning experience.

- S. F. Alaoui, K. Carlson, and T. Schiphorst. Choreography as mediated through compositional tools for movement: Constructing a historical perspective. In *Proceedings of the 2014 International Workshop on Movement and Computing*, pages 1–6, 2014.
- [2] O. Alemi, P. Pasquier, and C. Shaw. Mova: Interactive movement analytics platform. In *Proceedings of the 2014 International Workshop* on Movement and Computing, pages 37–42, 2014.
- [3] R. K. d. Anjos, C. Ribeiro, and C. Fernandes. Three-dimensional visualization of movement qualities in contemporary dance. In *Proceedings of the 5th International Conference on Movement and Computing*, pages 1–7, 2018.
- [4] A. Aristidou, D. Cohen-Or, J. K. Hodgins, Y. Chrysanthou, and A. Shamir. Deep motifs and motion signatures. ACM Transactions on Graphics (TOG), 37(6):1–13, 2018.
- [5] A. Behrends, S. Müller, and I. Dziobek. Moving in and out of synchrony: A concept for a new intervention fostering empathy through interactional movement and dance. *The Arts in Psychotherapy*, 39(2):107–116, 2012.
- [6] K. Carlson, P. Pasquier, H. H. Tsang, J. Phillips, T. Schiphorst, and T. Calvert. Cochoreo: A generative feature in idanceforms for creating novel keyframe animation for choreography. In *Proceedings of the 7th International Conference on Computational Creativity*, pages 380–387, 2016.
- [7] K. Carlson, T. Schiphorst, and C. Shaw. Actionplot: a visualization tool for contemporary dance analysis. In *Proceedings of the International Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging*, pages 113–120, 2011.
- [8] S. H. Chavoshi, B. De Baets, T. Neutens, G. De Tré, and N. Van de Weghe. Exploring dance movement data using sequence alignment methods. *PloS one*, 10(7):e0132452, 2015.
- [9] W. Forsythe and R. Sulcas. William forsythe: Improvisation technologies, 2000.
- [10] S. Hsueh, S. F. Alaoui, and W. E. Mackay. Understanding kinaesthetic creativity in dance. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2019.

- [11] M. Jacob and B. Magerko. Viewpoints ai. In Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition, pages 361– 362, 2015.
- [12] S. Jürgens, N. N. Correia, and R. Masu. Designing glitch procedures and visualisation workflows for markerless live motion capture of contemporary dance. In *Proceedings of the 7th International Conference* on Movement and Computing, pages 1–8, 2020.
- [13] M. Kapadia, I.-k. Chiang, T. Thomas, N. I. Badler, and J. T. Kider Jr. Efficient motion retrieval in large motion databases. In *Proceedings* of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games, pages 19–28, 2013.
- [14] M. Kyan, G. Sun, H. Li, L. Zhong, P. Muneesawang, N. Dong, B. Elder, and L. Guan. An approach to ballet dance training through ms kinect and visualization in a cave virtual reality environment. ACM Transactions on Intelligent Systems and Technology (TIST), 6(2):1– 37, 2015.
- [15] D. Long, L. Liu, S. Gujrania, C. Naomi, and B. Magerko. Visualizing improvisation in luminai, an ai partner for co-creative dance. In *Proceedings of the 7th International Conference on Movement and Computing*, pages 1–2, 2020.
- [16] K. E. Raheb, M. Stergiou, A. Katifori, and Y. Ioannidis. Dance interactive learning systems: A study on interaction workflow and teaching approaches. ACM Computing Surveys (CSUR), 52(3):1–37, 2019.
- [17] I. Rallis, N. Doulamis, A. Doulamis, A. Voulodimos, and V. Vescoukis. Spatio-temporal summarization of dance choreographies. *Computers & Graphics*, 73:88–101, 2018.
- [18] P. Subyen, D. Maranan, T. Schiphorst, and P. Pasquier. Mapping, meaning and motion: Designing abstract visualization of movement qualities. 2010.
- [19] W. D. Toolbox. Badco. and danielturing. Transmission in Motion: The Technologizing of Dance, page 118, 2016.
- [20] S. Whatley, R. E. K. Cisneros, and R. Gibson. Whole-body interaction learning for dance education: Deliverable 1.1 sota survey. 2016.
- [21] Q. Zhang and X. Guo. 3d human motion retrieval based on human hierarchical index structure. *Biology of sport*, 30(2):145, 2013.

### Multidimensional Animated Visualizations for Transgender Voice Training

Luna McNulty \*

**Brown University** 

#### ABSTRACT

We created nFormants, an application which visualizes vocal properties pertinant to gender voice training – including formant height, pitch, and resonance – synchronized with the playback of a user's voice recording. We asked users to compare three provided visualizations, and found that they considered our novel 2D visualization of pitch and resonance more useful than 1D pitch visualizations, as commonly found in similar software – compared to which many rated nFormants as superior. They also found ithe 2D visualization easier to understand than a full visualization of the formants, as found in complex speech analysis tools. Consequently, we believe that such visualizations have potential to facilitate the difficult process of vocal transition for transgender people.

Keywords: human-computer interaction, evaluation.

#### **1** INTRODUCTION

Transgender people report frequent distress from others misperceiving their gender because of their voice [3]. Many work with a speech-language pathologist to change their voice to better align with their gender identity. Such training typically focuses on altering the fundamental frequency (pitch) of the voice [9]. There is also substantial evidence for the effectiveness of resonance training [2], which seeks to alter the acoustic influence of the vocal tract by techniques such as raising the larynx.

Most transgender women<sup>1</sup> (62%) want to undergo formal voice therapy, but only 14% have done so [5], suggesting that many are left attempting to change their voice without the feedback of a professional. This can result in misdirected effort if such individuals focus on the wrong dimensions of their voices, such as by attempting to increase pitch while neglecting resonance. In doing so, they might also push certain dimensions beyond the appropriate range, to the point where it becomes difficult or unhealthy to maintain their voices in daily life. Existing software for gender voice visualizations likely exacerbates this problem by encouraging narrow focus on pitch, which is more easily measured than other pertinent vocal features. For example, a visualization from Ahmed and Borkin's Project Spectra presents users with a one-dimensional chart, plotting their pitch on a range from masculine to androgynous to feminine [1]. However, a user attempting voice feminization can easily reach the very top of the feminine range by speaking in a falsetto, even though this is unlikely to be read as a female voice.

We believe that such issues can be resolved by visualizing multiple dimensions of voice. We created the nFormants web application to test demonstrate this by visualizing formant heights and resonance in addition to pitch.

#### 2 THE VISUALIZATIONS

nFormants gives automated feedback to those undergoing gender voice training – who can submit a recording of their voice and view three different animated visualizations in sync with the playback.

#### 2.1 The Formant Tracker



Formants are bands of concentrated acoustic energy in a voice at the resonant frequencies above the fundamental (pitch). We follow the convention of notating pitch as  $f_0$ , with the first formant above it as  $f_1$ , the second as  $f_2$ , etc. Formant heights are substantially different between men and women [4] [11] [10] and have been shown to provide useful feedback for some transgender women undergoing voice training [6].

When the user submits a recording in nFormants, the backend produces a set of formant heights sampled at particular points in time using the open source phonetics program Praat with the default parameters. To obtain a smooth, continuous function suitable for animation, all displayed data are run through the following median filter:

$$\mu_i(t) = M\left\{ p_i \left| p \in S \land |t - p_0| < r \right\} \right\}$$

- *M* is a function giving the median of set.
- p is a vector containing values for formants 0-n, i.e.  $[f_0, f_1, ..., f_n]$ , with the time in seconds represented as  $p_0$ .
- *S* is the set of all such points returned by Praat.
- *r* is the filter radius in seconds. We used *r* = 2 in the application.

The formant tracker plots each formant  $f_i(t) = \mu_{i+1}(t)$  as a line graph superimposed over a scatter plot of the raw data points. Each formant is scaled differently to fit the common values in human voices, and its values are shown beneath for the current playback time. This visualization provides a highly accurate and complete picture of a voice. However, it requires special expertise to interpret, particularly with respect to voice gender. Therefore, we predicted that most users would struggle to make use of it.

<sup>\*</sup>e-mail: lmcnulty@cs.brown.edu

<sup>&</sup>lt;sup>1</sup>Trans men are less likely to seek voice therapy, as testosterone treatment typically has a significant effect on the voice on its own, so most data on this topic concerns trans women. However, the number of trans men who require voice therapy is frequently underestimated and has been estimated around 24% [8].

#### 2.2 The Pitch Tracker



The design of the pitch-tracker largely replicated that used by Project Spectra [1], plotting the pitch,  $P(t) = \mu_1(t)$ , on a vertical bar colored to correspond to likely gender perception with tick-marks for musical notes shown to the side.

Pitch trackers are the most widely-used type of visualization in current software for gender voice training, largely due to their simplicity and ease of implementation. Therefore, we predicted that users would find this visualization easy to use but not particularly reflective of voice gender perceptions, for the reasons outlined in the introduction.

#### 2.3 The Pitch/Resonance Graph



The Pitch/Resonance graph adds the second dimension of "resonance" to the pitch tracker. Resonance is an abstraction over the formant heights, representing the combined impact of the effective size of the passages that the voice travels through, including the laryngeal cavity, the pharynx, and the mouth. We obtain a single value for resonance with a weighted average of the formant heights:

$$R(t) = \sum_{i=1}^{n} w_i \mu_i(t)$$

- *n* is a constant for the number of formants included in the resonance calculation. We used *n* = 3 in the application.
- *w* is a constant vector containing weights for each formant. We used [0,0.5,0.5] in the application.

As the resonance increases, the point representing the user's voice moves to the warm-colored right side of the graph. As the pitch increases, the point moves up to the light-colored top half of the graph such that highly feminine voices will be represented in light-pink and highly masculine voices in dark blue. We predicted that this visualization would be substantially easier to understand

than the formant tracker and more useful and reflective of voice gender than the pitch tracker due to its relative simplicity.

#### **3** SURVEY & CONCLUSIONS

We made nFormants accessible online with a linked survey, and promoted it in the r/transvoice subreddit [7] and queer community spaces at Brown. The survey received 23 responses. Of the respondents, 21 (91.3%) were trying to feminize their voice, 1 (4.3%) was trying to masculinize his voice, and 2 (8.7%) had other goals. All reported a gender identity different from their assigned sex at birth. 9 (39.1%) had been training their voice for over a year, 4 (14.4%) between between 6 months and 1 year, and 10 (43.4%) for less than 6 months.

Users were asked to evaluate each visualization in the following respect, by agreement or disagreement with the following statements, with results in Table 1:

- "The visualization is easy to understand."
- "A visualization like this would be **helpful** in attaining my goals for my voice"
- "The data displayed reflect how I perceive voice gender."

	Easy	Helpful	Reflect
Formant Tracker	9 (39%)	15 (65%)	3 (13%)
Pitch Tracker	19 (83%)	15 (65%)	13 (57%)
Pitch/Resonance	16 (70%)	20 (87%)	12 (52%)

Table 1: Number of respondants who agreed or strongly agreed with the given statement for each visualization

Consistent with our expectations, a great majority of participants agreed that a visualization like the Pitch/Resonance graph would be helpful in attaining their goals for their voice, 17 (74%) saying that they "strongly agreed." This compares favorably against the pitch and formants tracker, which only 15 (65%) respondents each agreed would be helpful in attaining their goals.

Likewise, users found the pitch/resonance graph almost as easy to understand as the pitch tracker, which is a strong showing considering the already widespread use of pitch-tracking software in this domain. Both were found substantially easier to understand than the formant tracker which most users did not have the necessary knowledge to interpret.

The high ratings in helpfulness and ease of understanding suggest that the pitch/resonance graph is highly effective as a visualization. However, contrary to our expectations, users were no more likely to say that the pitch tracker and the pitch/resonance graph reflected their perceptions of voice gender than the pitch tracker. We believe that this is because several users reported inaccuracy in the resonance detection, particularly in relation to noise, which lowered overall scores. However, users were generally enthusiastic about the addition of resonance as a dimension, with one respondent wrote that "the resonance graph is especially helpful since it is more difficult to locate resonance visualizers." and seemed eager to use such a visualization once errors were resolved.

Users were additionally asked to rate the visualizations provided by nFormants against other software they had used for voice training purposes. 36% of respondents who had previously used software for voice training said that nFormants was better than anything they had used before. Of the apps/programs named by at least two users, nFormants was rated higher than all but Friture and Informant, both of which provide complex visualizations targeted towards export users. This suggests that multidimensional visualizations represent a substantial improvement in the software available for non-experts users undergoing voice training.

- [1] A. A. Ahmed and M. A. Borkin. Data visualization for transgender voice training.
- [2] L. Carew, G. Dacakis, and J. Oates. The effectiveness of oral resonance therapy on the perception of femininity of voice in male-tofemale transsexuals. *Journal of voice*, 21(5):591–603, 2007.
- [3] G. Dacakis, S. Davies, J. M. Oates, J. M. Douglas, and J. R. Johnston. Development and preliminary evaluation of the transsexual voice questionnaire for male-to-female transsexuals. *Journal of Voice*, 27(3):312–320, 2013.
- [4] S. J. Gallena, B. Stickels, and E. Stickels. Gender Perception After Raising Vowel Fundamental and Formant Frequencies: Considerations for Oral Resonance Research. *Journal of Voice*, 32(5):592–601, Sept. 2018.
- [5] S. James, J. Herman, S. Rankin, M. Keisling, L. Mottet, and M. Anafi. The report of the 2015 us transgender survey. 2016.
- [6] D. Kawitzky and T. McAllister. The Effect of Formant Biofeedback on the Feminization of Voice in Transgender Women. *Journal of Voice*, 34(1):53–67, Jan. 2020.
- [7] T. Kowalchuk. What trans\* folks talk about when they talk about voice: Learning about voice feminization on reddit. 2020.
- [8] U. Nygren, A. Nordenskjöld, S. Arver, and M. Södersten. Effects on voice fundamental frequency and satisfaction with voice in trans men during testosterone treatment—a longitudinal study. *Journal of Voice*, 30(6):766–e23, 2016.
- [9] J. Oates and G. Dacakis. Transgender voice and communication: Research evidence underpinning voice intervention for male-to-female transsexual women. *Perspectives on Voice and Voice Disorders*, 25(2):48–58, 2015.
- [10] A. P. Simpson. Phonetic differences between male and female speech: Phonetic differences between male and female speech. *Language and Linguistics Compass*, 3(2):621–640, Mar. 2009.
- [11] V. G. Skuk and S. R. Schweinberger. Influences of Fundamental Frequency, Formant Frequencies, Aperiodicity, and Spectrum Level on the Perception of Voice Gender. *Journal of Speech, Language, and Hearing Research*, 57(1):285–296, Feb. 2014.

## Visualizing Alzheimer's Disease Brain Pathology Using Deep Conditional Population Templates

Shuyan Wang\*

#### Brown University



Figure 1: Alzheimer's Disease Synthesis Tool Demo

#### ABSTRACT

In this work, we explore the current state of brain template generation and medical image visualization tools for Alzheimer's Disease. We identify the gap in current research and build a new visualization tool for exploring Alzheimer's Disease brains templates conditioning on different demographic variables. We train a convolutional neural network that outputs brain templates representing population-average data and propose an interface for the synthesized visualization. This tool will allow Alzheimer's Disease researchers and physicians to interactively explore brain anatomy as a function of user-controlled demographic variables such as age, sex, and disease stage, and uncover new insights that are not otherwise possible with individual scans.

**Keywords:** Alzheimer's Disease, deformable templates, brain atlases, MRI visualization.

#### **1** INTRODUCTION

Alzheimer's disease (AD) is a progressive disease that begins with mild cognitive impairment and eventually destroys memory and other cognitive functions. It is the most common form of dementia, affecting over 5 million Americans. While cognitive changes in AD are linked to brain changes that are observable with magnetic resonance imaging (MRI) and positron emission tomography (PET), the specific effects and patterns are often subtle and only apparent when looking across many individuals or across long periods of time. Such characteristics of AD made developing brain templates and a synthesis tool representing population-average data especially crucial for AD research and diagnosis.

We explore ways to interactively visualize brain changes associated with Alzheimer's Disease that reflect population-level features. The project is divided into two parts: synthesis and visualization. In the synthesis stage, we build a convolutional neural network (CNN) model using the framework of VoxelMorph [4]. We extend the conditional template framework developed in VoxelMorph to support new imaging modalities and conditional variables relevant to Alzheimer's disease. We train both the general template and conditional templates conditioning on patients' demographic variables such as age, sex and disease stage. In the visualization stage, we explore both 2D and 3D medical image visualization tools, experiment with Quantitative Imaging Toolkit [6], and sketch out a prototype for our AD synthesis visualization tool.

Although we did not finish the visualization tool this semester, our work and preliminary feedback point out a clear and optimistic future direction. The end tool will allow AD researchers to interactively explore brain anatomy as a function of user-controlled demographic variables and potentially discover new patterns or biomarkers that are not otherwise visible on individual scans. At the same time, physicians can compare between two synthesized cases or between a synthesized case and a real patient and help with diagnosis and treatment.

#### 2 SYNTHESIS

#### 2.1 The Data

The project uses 632 patients' structural MRI (T1-weighted) and diffusion MRI data from ADNI (Alzheimer's Disease Neuroimaging Initiative [1]). Patients not only vary by age and sex, but are also in different disease stages: from cognitive normal to early

<sup>\*</sup>e-mail: shuyan\_wang@brown.edu

mild cognitive impairment to late mild cognitive impairment to the AD stage. As a result, the dataset provides a good amount of information for us to train the templates on. All scans are preprocessed with standard steps, including affine spatial normalization and anatomical segmentations. Final images are cropped to  $160 \times 192 \times 224$ .

#### 2.2 The Model

Building deformable templates plays a fundamental role in many medical image analysis and computational anatomy tasks. However, it is often a challenge to crunch datasets of patients' medical images and capture structural relationships that may indicate the progression of diseases due to high structural variability and complexity across patients and scans.

To tackle this challenge, we first study how deformable registration and template generation work, and look at different methods and frameworks for generating templates. We then evaluate existing template generation frameworks based the goal of our project and the specific medical images data we are using. We decide to use VoxelMorph [4], which is a probabilistic model that learns a parametrized registration function from a collection of volumes. It is more computationally efficient compared to other conventional methods that uses an iterative process of template estimation and alignment, or diffeomorphic methods, such as [2] [3]. VoxelMorph computes on-demand conditional deformable templates by leveraging the entire image collection; the model also enables the use of multiple attributes to condition the template on, without needing to apply arbitrary thresholds or subdividing a dataset. The following figure shows the CNN model architecture of VoxelMorph.



Figure 2: CNN Model Structure. Each rectangle represents a 3D volume, generated using a 3D convolutional network layer. Arrows represent skip connections, which concatenate encoder and decoder features.

We write scripts and .csv files containing the conditional demographic variables to train the templates on. We train the general template and each conditional template for sex, age, and disease stage for 1500 epochs on a Lambda GPU machine.

#### **3** VISUALIZATION

We start the visualization stage by thinking about what forms of medical image visualizations can benefit the two target user groups, AD researchers and physicians, the most. Besides looking at 2D images slices, which most of existing brain atlas do, we further visualize the brain in 3D, helping users see brain atrophy that might not be clear from 2D slices where the changes may be isolated to certain lobes. Furthermore, the 3D approach can be extended to look at other structures, such as the ventricles and hippocampus, which are crucial in the research of Alzheimer's Disease.

#### 3.1 Visualization Design

With these goals in mind, we explore the Quantitative Imaging Toolkit [6], a software for interactive 3D visualization, processing and analysis of neuroimaging datasets developed by the collaborator of this research. We test out different views and functionalities of the tool and decide on the combination that works best for the purpose of AD synthesis. We design a comparison view that incorporates both 3D rotation and slicing capabilities with three 2D views lined up at the bottom. The interface makes it convenient for AD researchers to look closely at the synthesized scans for customized inputs, and for physicians to compare two scans side to side in one setting. This results in the interface demo shown in Figure 1.

#### 3.2 Visualization Evaluation

Due to the time constraint of this project, we have not yet deployed the tool to full functionality. However, once the visualization tool becomes fully functional, we plan to evaluate the tool based on two metrics. First, the accuracy of synthesized brain image generated, if it truly reflects the sub-population average, and secondly, the tool's effectiveness in real use cases of research and diagnosis, based on feedback from research scientists and physicians. To get an idea of whether we are headed to the right direction, we have collected some preliminary feedback through interviewing physicians at the Brown University Medical School Department of Neurology and showing them the demo of the interface. The feedback we received points out the novelty of such AD synthesis tool, as well as the usefulness of such view in surgery planning process that could be extended outside the scope of AD.

#### 4 RELATED WORK

Raman et al. [9] and Sun et al. [11] developed brain atlases representing the cortical regions that are routinely sampled at autopsy for AD diagnosis, which is the close to our synthesis work. On the visualization side, past approaches of brain pathology visualization usually focus on one specific biomarker, chemical or part of the brain [7] [8] [10]. A recent work presents an interactive tool to derive scan-specific parcellations from established atlases [5]. Our work is novel both in the specificity of Alzheimer's Disease research and in the generality of the process (building the template all the way to presenting the result to users) that could be applied to other diseases and research fields.

#### 5 CONCLUSION

In this work, we explore the current methods and tools for brain template generation and medical image visualization. We contribute our novel approach and visualization design to a specific problem at the intersection of Alzheimer's Disease research and visual analytics in neuroimaging. In terms of scientific contribution, the project opens up new knowledge in Alzheimer's Disease by training conditional brain templates on real patients' medical data, and synthesizing AD brain pathology. In terms of visualization contribution, the project proposes a new presentation of information and visualization catered specifically to brain template generation and analysis. Admittedly, the contributions are hard to be proved at this stage because we lack comprehensive feedback for the tool from actual users; we have not yet tested the AD brain templates because we do not have the subject matter expertise. However, we are confident that we successfully identified a meaningful gap in current research and made attempts to address the challenge. Our work laid solid ground work for future augmentation of training dataset and iterative development of the Alzheimer's Disease synthesis tool.

#### ACKNOWLEDGEMENTS

This work is for CS2370 at Brown University. The author would like to thank collaborator Ryan P. Cabeen at University of Southern California for generous guidance and support throughout the project. The author would also like to thank David H. Laidlaw and Fumeng Yang for directions of this project.

## Visualization Methods to Compare Metrics to Quantify Forest Structure

**Christopher Wong\*** 

**Brown University** 

#### ABSTRACT

We present a novel visualization tool to allow ecologists to more easily explore Light Detection and Ranging (Lidar) forest data and quantify forest structure. This tool builds on existing visualization tools as it has the ability to integrate structure summary metrics with raw lidar data that allows ecologists to gain an intuitive sense of how the metrics relate to their own sense of what the forest looks like. When involving our collaborators, we assessed that researchers found the integration of the data more informative than standard depictions of lidar data. We will also present a review of three software packages for visualizing point cloud data and the summary information.

Keywords: Lidar, point clouds

#### **1** INTRODUCTION

We investigate the optimal software and integration methods for displaying point cloud lidar data. Lidar stands for light detection and ranging and works by emitting light pulses at an area of interest and records their return time. It has become widely used in the study of forests and the monitoring of a forest's health given its ability to survey large areas quickly while capturing a high level of detail.

Researchers leverage these huge datasets to both visualize large swaths of land as well generating models to predict different forest metrics related to health [2]. This project attempts to bridge the gap between the modeling and the visualization by creating a tool that can incorporate summary metrics into the lidar data so that the results can be validated by researchers. Dechesne et al. [1] support the need for innovative tools to draw summary information from lidar data, including machine learning. This project displays these summary stats in a way that scientists can compare and classify themselves as well as toggle through different metrics and compare. David L. Kao et al. [4] have used visualizations to display structure characteristic distributions. Our project takes similar summary statistics and integrate them with the raw lidar data in one visualization tool. Fujisaki et al. [3] supports the exploration of user studies on what visualization techniques are most appropriate and useful for ecologists in understanding forest structure.

#### 2 SOFTWARE REVIEW

In the process of developing the final visualization tool, we explored a range of software packages, namely Paraview, Unity, and OpenGL. Each of these packages were used with the goals being 1) plotting point cloud data, in the form of X,Y,Z coordinates, in world space that can be viewed and explored by users, 2) calculate and integrate different forest structure summary metrics, and 3) provide a user interface that allows for changing between metrics to facilitate exploratory analysis. We examined height and spatial density as our metrics. Our original plan was to explore different methods

\*e-mail: christopher\_wong2@cs.brown.edu

of integration, but due to time constraints we ended up limiting this search to just using color mapping.

#### 2.1 Paraview



Figure 1: Paraview visualization

Of the three software packages, Paraview provided the most support and ease towards accomplishing the first goal, plotting the point cloud data. Paraview's interface allows for a simple upload of a CSV containing the point cloud coordinates. This produced the above visualization. The severe limitation of Paraview, however, is that it can only color points with a solid color, which made our goal of color mapping the data impossible. Moreover, Paraviews own interface is robust, but also rigid and thus we also could not create our own user interface

#### 2.2 Unity



Figure 2: Unity visualization with UI and color mapping according to height

The Unity Game Engine proved to be a much better option in accomplishing our two goals. The relative difficulty of the two tasks was the same as they were with Paraview, however for different reasons. Converting the point cloud coordinates required generating meshes first and rendering the meshes into the game engine. We were able to use an open source project which handled the basics of the mesh functionality [5], however we had to adapt it slightly for our specifications. For our second goal, the color mapping itself was simple, with each points color being a settable attribute of the meshes, but the calculation of the various summary metrics proved to be harder. This, however, is not directly related to the software package and would persist across any implementation. Unity provides many options for creating a UI while remaining flexible enough for our purposes.

One drawback of the Unity implementation is that the extra step of generating meshes and rendering causes a performance slowdown. The forest lidar dataset we used contains 10,001,666 points, which when rendering from scratch, takes around a minute. We were able to store intermediate mesh data upon a first render that cuts down on rerender time.



Figure 3: The flow in Unity implementation. Upon rerender, the program skips the first and second step, which cuts down on render time considerably

#### 2.3 OpenGL



The performance drawback of Unity prompted us to try one more software package, OpenGL. We suspected that a low level program that can interact directly with GPUs could help with performance speed. We were also able to use an existing interface that we repurposed from CS1230 Computer Graphics, a course at Brown, that accomplished goal 3. We accomplished goal 1 using vertex buffer objects (VBO) and vertex array objects (VAO). Our hope that OpenGL would provide a performance boost was proven false at this point as our tool experienced lag that made it unusable when rendering just a few hundred thousand of the total 10 million points. This discovery made OpenGL a non-starter and goal 2 irrelevant. While performance could no doubt be improved given certain optimizations, we decided this was not a worthwhile use of time.

#### **3** DISCUSSION

With only the Unity tool accomplishing our three goals, I met with collaborators and gained some initial strengths and weaknesses of the current implementation. The visualization displayed the point cloud data in a way that effectively captured the visual structure of the forest. When certain summary metrics were color mapped, however, this structure was somewhat obscured. The interactivity provided by the UI helped mitigate this loss of information as toggling between metrics allowed direct comparison. This was further alleviated by rendering multiple point clouds adjacent to each other simultaneously, while color mapped with different metrics.

Another functionality we implemented was the ability for users to adjust the size of voxel, the subarea within which we calculate certain summary metrics. The left image in Figure 4 is a mapping of spatial density where density is a calculation of the number of points located in each of approximately 300,000 equally sized voxels dividing up the whole point cloud.



Figure 4: Left: Spatial Density, Right: Height. The structure of the forest was better visualized in the height metric, while spatial density obscured it slightly

#### 4 CONCLUSION

We produced a novel visualization tool that builds upon existing tools by allowing for color mapping of summary metrics and increased user control over the display and calculation of said metrics. When meeting with collaborators, we determined that the interactivity coupled with the color mapping were effective in integrating and understanding summary metrics in tandem with the raw representation of lidar data. We also explored the strengths and weaknesses of three different software packages, Paraview, Unity, and OpenGL, for visualizing point cloud data, incorporating summary metics through color mapping and building a UI to facilitate user control of the tool.

- [1] C. Dechesne, C. Mallet, A. Le Bris, V. Gouet, and A. Hervieu. Forest Stand Segmentation Using Airborne LIDAR Data and Very High Resolution Multispectral Imagery. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 49B3:207–214, June 2016.
- [2] T. L. Erdody and L. M. Moskal. Fusion of lidar and imagery for estimating forest canopy fuels. *Remote Sensing of Environment*, 114(4):725 – 737, 2010.
- [3] I. Fujisaki, D. L. Evans, R. J. Moorhead, D. W. Irby, M. J. Mohammadi-Aragh, S. D. Roberts, and P. D. Gerard. Stand Assessment through Lidar-Based Forest Visualization Using Immersive Virtual Environment Technology. *Forest Science*, 54(1):1–7, 02 2008.
- [4] D. L. Kao, M. G. Kramer, A. L. Love, J. L. Dungan, and A. T. Pang. Visualizing distributions from multi-return lidar data to understand forest structure. *The Cartographic Journal*, 42(1):35–47, 2005.
- [5] G. LLorach. Unity point cloud free viewer. GitHub, July 2014.

- [1] Alzheimer's disease neuroimaging initiative.
- [2] J. Ashburner. A fast diffeomorphic image registration algorithm. *Neuroimage*, 38(1):95–113, 2007.
- [3] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical image analysis*, 12(1):26–41, 2008.
- [4] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca. Voxelmorph: a learning framework for deformable medical image registration. *IEEE transactions on medical imaging*, 38(8):1788–1800, 2019.
- [5] R. G. Bayrak, N. Hoang, C. B. Hansen, C. Chang, and M. Berger. Pragma: Interactively constructing functional brain parcellations. *arXiv preprint arXiv:2009.01697*, 2020.
- [6] R. Cabeen, D. Laidlaw, and A. Toga. Quantitative imaging toolkit: software for interactive 3d visualization, data exploration, and computational analysis of neuroimaging datasets. *ISMRM-ESMRMB Ab*stracts, pages 12–14, 2018.
- [7] V. Kepe, S.-C. Huang, G. W. Small, N. Satyamurthy, and J. R. Barrio. Visualizing pathology deposits in the living brain of patients with alzheimer's disease. *Methods in enzymology*, 412:144–160, 2006.
- [8] A. K. L. Liu, M. E. Hurry, O. T.-W. Ng, J. DeFelice, H. M. Lai, R. Pearce, G. T.-C. Wong, R. C.-C. Chang, and S. M. Gentleman. Bringing clarity to the human brain: visualization of lewy pathology in three dimensions. *Neuropathology and applied neurobiology*, 42(6):573–587, 2016.
- [9] M. R. Raman, C. G. Schwarz, M. E. Murray, V. J. Lowe, D. W. Dickson, C. R. Jack Jr, and K. Kantarci. An mri-based atlas for correlation of imaging and pathologic findings in alzheimer's disease. *Journal of Neuroimaging*, 26(3):264–268, 2016.
- [10] R. F. Roberts, R. Wade-Martins, and J. Alegre-Abarrategui. Direct visualization of alpha-synuclein oligomers reveals previously undetected pathology in parkinson's disease brain. *Brain*, 138(6):1642– 1657, 2015.
- [11] W. Sun, Y. Tang, Y. Qiao, X. Ge, M. Mather, J. M. Ringman, Y. Shi, A. D. N. Initiative, et al. A probabilistic atlas of locus coeruleus pathways to transentorhinal cortex for connectome imaging in alzheimer's disease. *NeuroImage*, 223:117301, 2020.