A Hybrid Approach for Image Vectorization for Semi-Geometric Images

Jonathan Lam, Derek Lee, Victor Zhang

Prof. Sam Keene
The Cooper Union for the Advancement of Science and Art

Abstract

Vector (shape-based) images are a useful image representation that may be more efficient than traditional raster (pixel-besed) formats. Our project aims to develop an image vectorization method (a tool to convert from raster to vector format) specialized towards geometric images, combining the benefits of edge tracing and sampling methods.

Proposed Methods

A raster image is first sampled using blue noise sampling (BNS) [4], which generates a point cloud with variable density. This point cloud is simplified using the quadric error metric (QEM) [2]. We also run the image through multiple color scans using the Potrace edge tracing algorithm [3]. We improve edges in the mesh using the Potrace paths and edge detection methods, and then triangulate the point cloud. We evaluate the accuracy of the generated image using MSE, file size, and content loss [1].

Conclusions

We have implemented a basic framework for vectorizing raster images, primarily based on blue-noise sampling [4] and Potrace [3]. Our proposed framework involves performing blue-noise sampling on an image, merging sampled points with the Potrace output, triangulation, and exporting to an SVG file. While this method currently does not generate a highly efficient representation, it performs better than BNS and Potrace on some metrics, particularly accuracy. Our method is able to both handle gradient patches better than Potrace, and represent edges better than blue-noise sampling. There is much future work remaining to further tune this model for efficiency.

References

[1] Vincent Dumoulin, Jonathon Shlens, and Manjunath Kudlur. *A Learned Representation For Artistic Style*. 2017. arXiv: 1610.07629 [cs.CV].

[2] Hugues Hoppe. "New quadric metric for simplifying meshes with appearance attributes". In: *Proceedings Visualization'99 (Cat. No. 99CB37067)*. IEEE. 1999, pp. 59–510.

[3] Peter Selinger. "Potrace: a polygon-based tracing algorithm". In: *Potrace (online), http://potrace. sourceforge. net/potrace. pdf (2009-07-01)* 2 (2003).

[4] Jiaojiao Zhao, Jie Feng, and Bingfeng Zhou. "Image vectorization using blue-noise sampling". In: *Imaging and Printing in a Web 2.0 World IV*. Vol. 8664. International Society for Optics and Photonics. 2013, 86640H.

Results

Figure 1: Original image

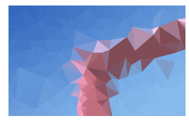


Figure 2: BNS image

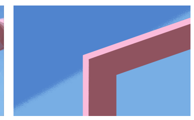


Figure 3: Potrace image

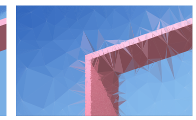


Figure 4: Hybrid image

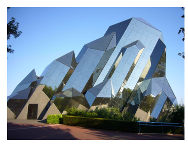


Figure 5: Original image

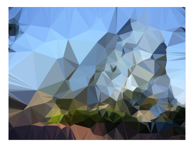


Figure 6: BNS image

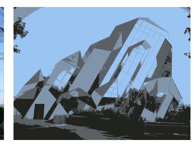


Figure 7: Potrace image

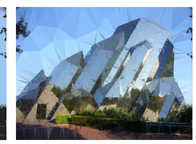


Figure 8: Hybrid image

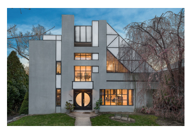


Figure 9: Original image

Figure 10: BNS image

Figure 11: Potrace image

Figure 12: Hybrid image

Figure 13: Original image

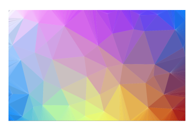


Figure 14: BNS image

Figure 15: Potrace image

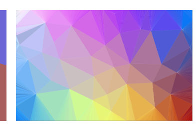
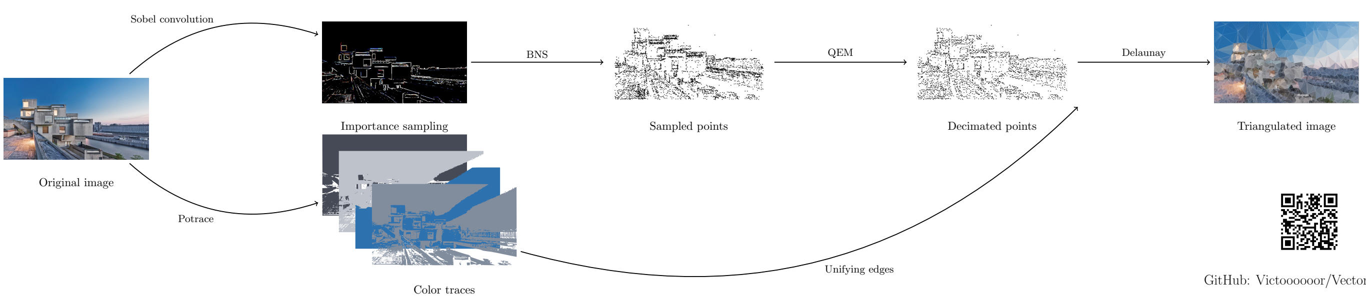


Figure 16: Hybrid image

Diagram of Proposed Methods



GitHub: Victoooooor/Vectorize-Arch

LaTeX TikZposter

Image Vectorization for Architecture

Jonathan Lam, Derek Lee, Victor Zhang

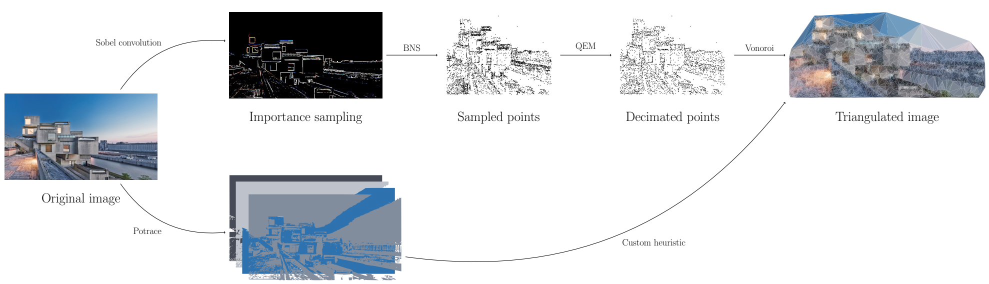
Prof. Samuel Keene
The Cooper Union for the Advancement of Science and Art

Abstract

Vector (shape-based) images are a useful image representation that may be more efficient than traditional raster (pixel-besed) formats. Our project aims to develop a image vectorization method (a tool to convert from raster to vector format) specialized towards architecture images (or other highly-geometric images), and explores the potential of vector-based images in the architecture design process and as machine learning preprocessing.

Algorithm/Architecture

A raster image is first sampled using blue noise sampling (BNS) [4], which generates a point cloud with variable density. This point cloud is simplified using the quadric error metric (QEM) [2]. We also run the image through multiple color scans using the Potrace edge tracing algorithm [3]. Then the point cloud is vectorized using a triangulation method and the edges are improved using the Potrace scans, and is outputted to a SVG format. We evaluate the accuracy of the generated image using content loss [1].



Sample Images

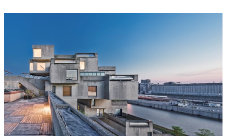


Fig. 1: Original image

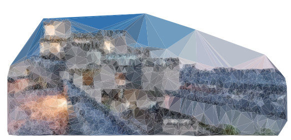


Fig. 2: Vectorized image

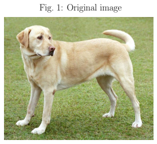


Fig. 3: Original image

Fig. 4: Vectorized image

Future Work

For the next iteration of our project, we need to combine our sampled points with the Potrace algorithm, which should improve the quality of edges on the final result. This will likely involve heuristic methods that directs mesh vectorization using the Potrace curves. The end goal is to improve the heuristics well enough to produce edges that resemble architectural line drawings.

References

[1] Vincent Dumoulin, Jonathon Shlens, and Manjunath Kudlur. *A Learned Representation For Artistic Style*. 2017. arXiv: `1610.07629 [cs.CV]`.

[2] Hugues Hoppe. "New quadric metric for simplifying meshes with appearance attributes". In: *Proceedings Visualization'99 (Cat. No. 99CB37067)*. IEEE. 1999, pp. 59–510.

[3] Peter Selinger. "Potrace: a polygon-based tracing algorithm". In: *Potrace (online), http://potrace.sourceforge. net/potrace. pdf (2009-07-01)* 2 (2003).

[4] Jiaojiao Zhao, Jie Feng, and Bingfeng Zhou. "Image vectorization using blue-noise sampling". In: *Imaging and Printing in a Web 2.0 World IV*. Vol. 8664. International Society for Optics and Photonics. 2013, 86640H.

GitHub: Victoooooor/Vectorize-Arch

Design of Phase-Locked Loop

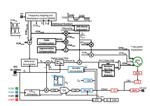
Evan Goldstein, Sophie Jaro, Hadassah Yanofsky

Problem Statement

A Phase Locked Loop (PLL) is a system that generates an output signal that has a constant phase difference from a reference input signal

These are commonly used for clock signal synthesis and can be used to multiply a reference clock by very precise amounts.

Our specific system will be used for bluetooth LNE applications (2.4 GHz)



C. -C. Li, M. -S. Yuan, C. -C. Liao, Y. -T. Lin, C. -H. Chang and R. B. Staszewski, "All-Digital PLL for Bluetooth Low Energy Using 32.768-kHz Reference Clock and ≤0.45-V Supply," in *IEEE Journal of Solid-State Circuits*, vol. 53, no. 12, pp. 3660-3671, Dec. 2018, doi: 10.1109/JSSC.2018.2871632.

Background Research

We used RF Microelectronics 2nd Ed. by Razavi for the bulk of our research into this topic.

There are a bunch of papers that we are using as reference but they have complex circuits aren't useful for industry.

Although other PLLs exist, ours will improve on those by consuming low power while maintaining clean output signal (low noise).

Table 1. PLL Performance Metrics for Bluetooth LE Application in Literature

Ref, Year	Frequency Range	Power Consumption	Phase Noise	Fractional Spur
[4], 2016	2.05-2.55GHz	1.4mW	-92 dBc/Hz @5MHz	-60 dBc
[5], 2017	1.8-2.5GHz	0.67mW	-103 dBc/Hz @1MHz	-56 dBc
[6], 2021	1.8 - 2.7 GHz	2.76mW	-112 dBc/Hz @10MHz	44 dBc

[4]"A Fully Integrated Bluetooth Low-Energy Transmitter in 28 Nm CMOS With 36% System Efficiency at 3 DBm." *IEEE Xplore Full-Text PDF:* https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7476842.
[5]"24.7 a 673μw 1.8-To-2.5GHz Dividerless Fractional-N Digital PLL with an Inherent Frequency-Capture Capability and a Phase-Dithering Spur Mitigation for IOT Applications." *IEEE Xplore*, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7870440
[6]"An Injection-Locked Ring-Oscillator-Based Fractional-N Digital PLL Supporting Ble Frequency Modulation." *IEEE Xplore*, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9737302.

Proposed Solutions

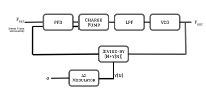
We have various parts that we wanted to work on improving.

We also wanted to work on improving phase noise and power consumption.

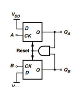
We worked on virtuoso, verilog and simulink to build and test our individual circuits.

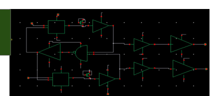
Experimental Solution

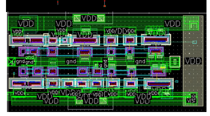
We split up the system into various subblocks, then combined all the individual components into the complete system and manage the interconnection.



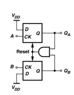
PFD

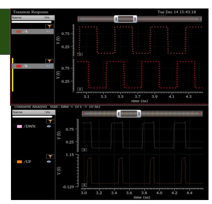




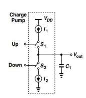


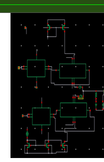
PFD Results

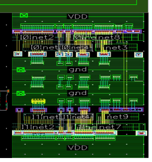


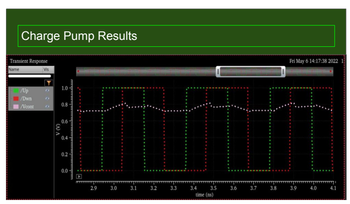


Charge Pump

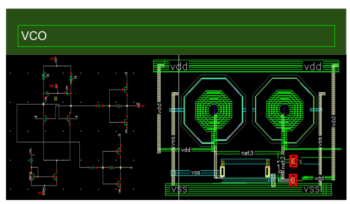




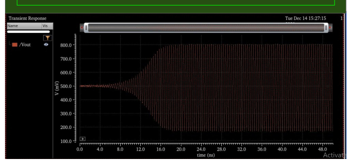




Charge Pump Results
Transient Response
Fri May 6 14:17:38 2022
Name
Vis
/Up
/Dwn
/Vcont
1.0
0.8
0.6
0.4
0.2
0.0
V (V)
2.9
3.0
3.1
3.2
3.3
3.4
3.5
3.6
3.7
3.8
3.9
4.0
4.1
time (ns)

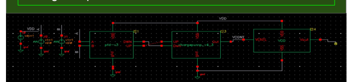

VCO
vdd
vdd
net3
vss
vss

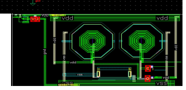
VCO Results



Transient Response
Tue Dec 14 15:27:15
Name
Vis
/Vout
800.0
700.0
600.0
500.0
400.0
300.0
200.0
100.0
V (mV)
0.0
4.0
8.0
12.0
16.0
20.0
24.0
28.0
32.0
36.0
40.0
44.0
48.0
time (ns)

Analog Components

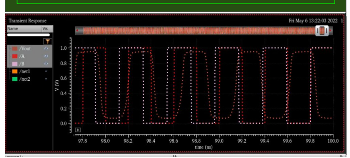




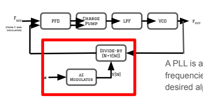
PFD, CP, VCO combined schematic

VCO layout dominates due to large inductors

Analog Components Results



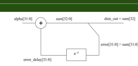
Fractional-N PLL

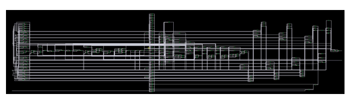


A PLL is able to synthesize fractional frequencies by dividing by N plus some desired alpha.

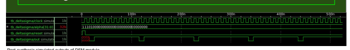
Alpha can be adjusted in tiny increments to achieve fine resolution frequency synthesis.

Delta Sigma Modulator

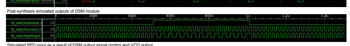




Delta Sigma Modulator Results

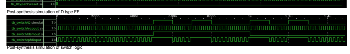


Post-synthesis simulated outputs of DSM module



Simulated PFD input as a result of DSM output signal control and VCO output

Post-synthesis simulation of D type FF



Post-synthesis simulation of switch logic

Next Steps

Next steps would be:

1. Combine the analog and digital components
2. Tapeout
3. Develop package and test procedures for chip

Initial plan was to import digital netlist into Cadence so we can combine with all analog circuits to run top simulation. But, we assumed we had the schematic view for each digital standard cells (we didn't – the TSMC digital library does not have schematic view). This week, we tried to extract the netlist of each analog block (VCO, PFD, and CP) and then build up test bench.

Thank You

Professor Koo

Professor Shlayan

Benchmarking Josephson Junctions for Scalable Quantum Computing

Tamar Bacalu, Alexa Jakob, Mark Koszykowski

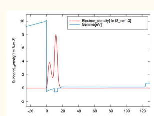
Objective

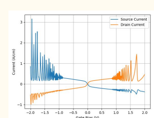
Verification of quantum devices is difficult to scale because of the time constraints associated with measurements. Our goal is to link characteristics of JJFETs at room & cryogenic temperatures to enable faster verification.

We do this by:

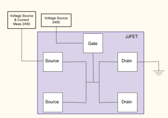
- Simulating devices at room temperature
- Measuring IV curves at room temperature
- Comparing R_N to resistance at room temperature
- Determining a mathematical relationship

Simulation Results

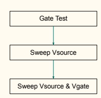




Measurement Setup

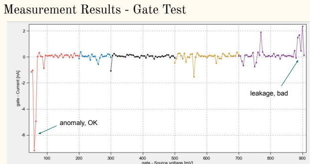


Four-point measurement circuit with voltage bias

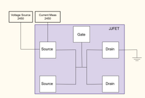


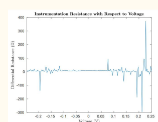
Measurement procedure

Measurement Results - Gate Test



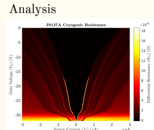
Results & Analysis



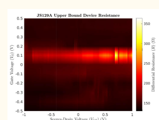


Average contact resistance: 6.3 Ohms

Analysis



Cryogenic Temperature



Room Temperature

Conclusion

- Difficult to draw conclusions: lack of data and devices, different measurement setups, yet some evidence to support cryogenic-high temperature relationship
- Provided documented framework for future students interested in project

Many Thanks!

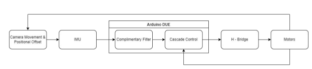
- Cooper Union: Neveen Shlayan
- NYU Shabani Lab: Javad Shabani, Mehdi Hatefipour, Billy Strickland, Mohammed Farzaneh
- NYU Tandon School of Engineering: Zhujun Huang


Cost Effective Gimbal
ARS GRATIA ARTIS
Anthony Belladonna
Philip Blumin
Paul Cucchiara
Daniel Kim

What is a Gimbal?

Features

- Rotational Stability in 2 dimensions
  - (Roll and Pitch)
- Custom MOSFET H - Bridges
- Custom Controller Shield



Gimbal Implementation

Hardware

- Microcontroller
- IMU
- DC Motors
    - Hall Encoders + Planetary Gearboxes
- H-Bridges
- Power Shield
- Power Supply
- Aluminum Frame

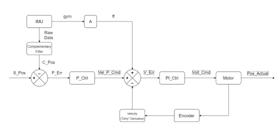
Sensor Fusion Algorithm

- Complementary Filter

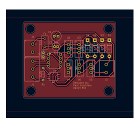
Feedback Controls

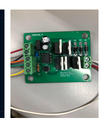
- Cascaded Controller

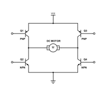
Cascade Control



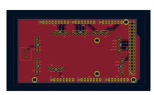
Custom H - Bridges

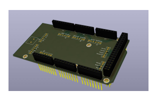




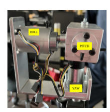


Custom Power Shield





Frame



Final Cost

3 Motors - $72

3 H-bridges - $42

Microcontroller - $40

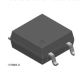
IMU - $3

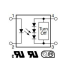
PCB - $2

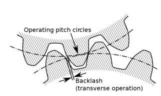
Frame - $5

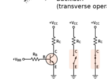
Difficulties & Lessons Learned

- Expensive motors are worth it
- Multiple IMUs are necessary
- Account for microcontroller output









Thank you!

A special thanks to:

Professor Shlayan

Sinisa Janjusevic

Michael Giglia

Look Ahead n Drive (LAnD)

Henry Son

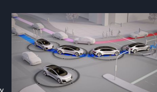

Intelligent Ground Vehicle Competition
IGVC
Intelligent Ground
Vehicle Competition

IGVC Rules

Sign / Obstacle	Dimensions
"Road Closed"	24" H x 30" W minimum height from ground is 5 feet
"One Way"	12" H x 36" W minimum height from ground is 5 feet
"Stop"	24" H x 24" H minimum height from ground is 5 feet

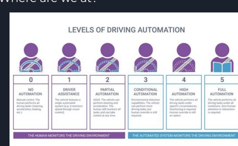
What’s the practical application?

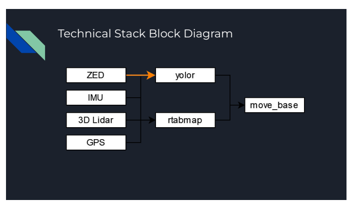
- Reliable and quicker transportation
- Reduce deaths from crash fatalities
- Possibly reduce congestion
- Increase productivity from saving time and energy actively driving

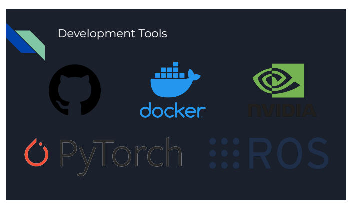


Source: https://drivemag.com/news/pros-and-cons-of-self-driving-cars-will-avs-solve-the-traffic-problem

Where are we at?

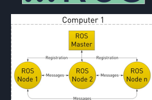



Technical Stack Block Diagram
ZED
IMU
3D Lidar
GPS
yolor
rtabmap
move_base

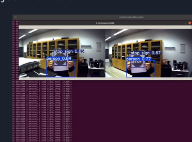

Development Tools
docker
NVIDIA
PyTorch
ROS

Object Detection

ROS



Object Detection

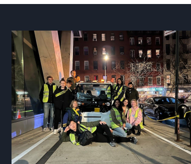


Future Work

- Create a publisher from the YOLOR node that gives the location of the center of mass of the object with respect to the camera and the object classification
- Optimize the memory usage for the algorithm
  - Shrinking the model
  - Write better code…

Thank you!

- Professor Shlayan
- Mike Giglia
- Dan Mezhiborsky
- Technical Team



Prediction of progression from CKD 4 to CKD 5

A machine learning approach

Presenter: Min Cheng

Group: Min Cheng in collaboration with
Dr. Lili Chan, Dr. Wonsuk Oh
from Mount Sinai

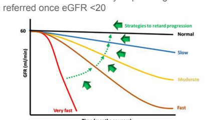
Chronic Kidney Disease stages

CKD stage[a]	GFR (ml/min/1.73m^2 body surface)	Description
1	≥90	Normal or increased GFR, with other evidence of kidney damage
2	60–89	Slight decrease in GFR, with other evidence of kidney damage
3a 3b	45–59 30–44	Moderate decrease in GFR, with or without other evidence of kidney damage
4	15–29	Severe decrease in GFR, with or without other evidence of kidney damage
5	<15	Established renal failure

CKD, chronic kidney disease; GFR, glomerular filtration rate; NICE, National Institute for Health and Care Excellence.
a. Use the suffix (p) to denote the presence of proteinuria when staging CKD.
Source: National Collaborating Centre for Chronic Conditions (2008).

Patients can progress at different rate

Early identification can allow for dialysis planning.
-- can be referred once eGFR <20



Kidney Failure Risk Calculator - Tangri

Kidney Failure Risk Equations

--Cohort
CKD stages 3 to 5
Mean eGFR: 36 mL/min/1.73 m_2

--Outcome:
Probability of kidney failure at 5 years
Dialysis
Transplantation

Table 2. Hazard Ratios and Goodness of Fit for Sequential Models in the Development Data Set[a]

Variable	Models 1	2	3	4	5	6	7
Baseline GFR, per 5 mL/min/1.73 m^2		0.54	0.57	0.58	0.60	0.61	0.64
Age, per 10 y	0.86	0.75	0.80	0.80	0.79	0.82	0.82
Male sex	1.03[b]	1.46	1.26	1.27	1.34	1.16[b]	1.26
Log spot urine ACR[c]			1.60	1.61	1.55	1.42	1.37
Diabetes				0.86[b]			0.88[b]
Hypertension				1.17[b]			0.88[b]
Systolic BP, per 10 mm Hg					1.15		1.14
Diastolic BP, per 10 mm Hg					1.10		1.15
Body weight, per 10 kg					0.91		0.91
Serum albumin, per 0.5 g/dL						0.84	0.83
Serum phosphate, per 1.0 mg/dL						1.27	1.34
Serum bicarbonate, per 1.0 mEq/L						0.92	0.93
Serum calcium, per mg/dL						0.81	0.82
C statistic[d]	0.56	0.89	0.91	0.91	0.92	0.92	0.92
Akaike Information Criterion[d]	5553	4834	4520	4521	4463	4432	4378
P value		<.001	<.001	.40	<.001	<.001	<.001

Abbreviations: ACR, albumin-to-creatinine ratio; BP, blood pressure; GFR, glomerular filtration rate.
[a] Data are presented as hazard ratios unless otherwise specified. Models 2, 3, and 6 columns indicate models based on laboratory data. *P* values are for comparison of C statistics between successive models, except for models 5, 6, and 7, which are compared with model 3.
[b] Hazard ratios with $P > .05$; all other hazard ratios are significant (ie, $P < .05$).
[c] Hazard ratio for ACR represents a 1.0 higher ACR on the natural log scale. For the average patient with 20 mg/g of albuminuria, this represents an increase to 55 mg/g.
[d] Null values for C statistic and Akaike Information Criterion are 0.50 and 5569, respectively. Higher values for C statistic and lower values for Akaike Information Criterion indicate better models.

Tangri, Navdeep & Inker, Lesley & Griffith, John & Tighiouart, Hocine & Djurdjev, Ognjenka & Naimark, David & Levin, Adeera & Levey, Andrew. (2011). A Predictive Model for Progression of Chronic Kidney Disease to Kidney Failure. JAMA : the journal of the American Medical Association. 305. 1553-9. 10.1001/jama.2011.451.

Problem Statement

- **Can we build a machine learning model using longitudinal EHR data to generate a risk prediction model for progression from CKD 4 to CKD5**?

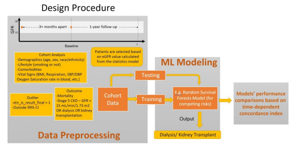
- **Hypothesis**
    1. The use of ML models can improve prediction accuracy compared to statistical models.
       --We all know that this hypothesis is no longer a hypothesis; rather, this can be a fact.

    2. The use of competing risk models can improve prediction accuracy than non-competing risk models.

    3. The use of annual laboratory data, e.g., Hb, eGFR, SCr, SCa, etc., can improve prediction accuracy than models using the latest known laboratory data.

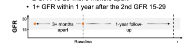
Flowchart Overview

Design Procedure

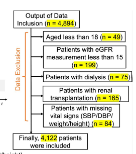


Methods and Results

- **Study design**
  - Mount Sinai Health System
  - Observation: 1/1/2011 – 12/31/2019
  - Inclusion
    - 2 GFR 15-29 with 3 months apart
    - 1+ GFR within 1 year after the 2nd GFR 15-29



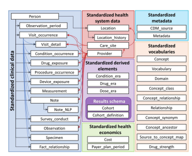
  - Exclusion
    - Aged less than 18
    - Patients with eGFR measurement less than 15
    - Patients with dialysis
    - Patients with renal transplantation
    - Patients with missing vital signs (SBP/DBP/weight/height)



Methods and Results

Data collection and Measurements

Mount Sinai Data Warehouse:
-- OMOP Common Data Model



Methods and Results

- **Data collection and Measurements**
  - Feature tables

person
person_id
birth_datetime
death_datetime
gender_concept_code
race_concept_id

observation
person_id
observation_date
observation_concept_id
value_as_concept_id

measurement
person_id
visit_occurrence_id
measurement_concept_id
measurement_date
value_as_number
unit_concept_id

condition_occurrence
person_id
xtn_epic_diagnosis_id
condition_start_date

  - Supplementary tables

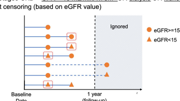
visit_occurrence
visit_occurrence_id
visit_concept_id

concept
concept_id
concept_code
concept_name

concept_relationship
concept_id_1
concept_id_2

Methods and Results

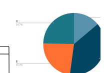
- **Data preprocessing**
  - Outcomes
    - Potential outcomes
      - Mortality
      - Stage5 CKD – GFR < 15 mL/min/1.73m² OR dialysis OR kidney transplantation
    - Right censoring (based on eGFR value)



Methods and Results

- **Data preprocessing**

– Demographics
• Age • Sex – male, female • Race/ethnicity – white, black, hispanic, other
– Complications
• Comorbidities used in Elixhauser comorbidity index
– Vital signs (outpatient setting)
• SBP – LOINC 8480-6 • DBP – LOINC 8462-4 • Pulse – SNOMED 78564009 • BMI – body weight (LOINC 8302-2) and body height (LOINC 29463-7) • Body temperature – LOINC 8310-5 • Respiratory rate – LOINC 9279-1 • Oxygen saturation in Blood – LOINC 20564-1
-- **Laboratory data** (GFR, serum creatinine, hemoglobin, serum calcium, serum phosphate, serum albumin, serum bicarbonate, urine albumin, urine creatinine)
-- Treatment



Outcome:
--Survival outcome
→features in the models
--Competing risk
outcome

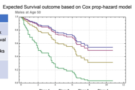
Ongoing and Future Works

- **Models**

	Statistical Models		ML Models	
	Adjusted	Adjusted+ Competing risk	Adjusted	Adjusted+ Competing risk
Semi-parametric	• Cox proportional hazard model	• Fine-Gray model	• Random survival forests • XGBoost • DeepSurv	• Random survival forests for competing risks
Parametric	• Weibull proportional ha zards model		• Deep recurrent survival analysis	• DeepHit

- **Performance metric**
  - Time-dependent concordance index

Expected Survival outcome based on Cox prop-hazard model



Many thanks to

Dr. Lili Chan (Nadkarni Lab, Icahn School of Medicine at Mount Sinai)

Dr. Wonsuk Oh (Nadkarni Lab, Icahn School of Medicine at Mount Sinai)

Prof. Neveen Shlayan (EE Department, Cooper Union)

Prof. Eric Lima (MechE Department, Cooper Union)

Exploring Feedback Control for Autonomous Vehicles

Nathaniel Kingsbury

Custom Motor Drive for Electric Racing Applications

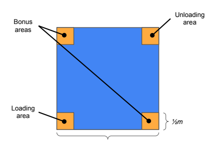
Samuel Shersher

H_2Go: Water-Powered Cart

Brandon Ho (ME'22), Jared Jacobowitz (ME'22), and
Josh Yoon (EE'22)
Advisors: Prof. David Wootton and Prof. Neveen Shlayan

Introduction

The Competition



H_2Go Competition Requirements

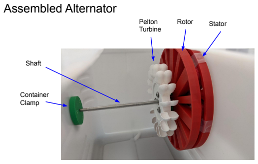
The device must fit within a 50cm x 50cm x 50cm box

The device must be remote-controlled

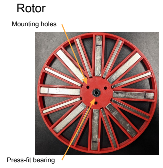
The device cannot have any initially stored energy

Design

Assembled Alternator

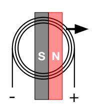


Rotor

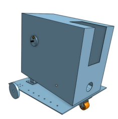

Mounting holes
Press-fit bearing

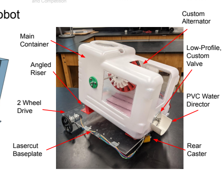

Strong-Field Side
Weak-Field Side

Stator

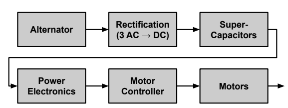


Our Competition Robot



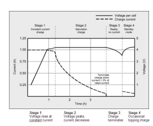


Electronics Overview



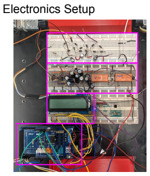
Why supercapacitors?

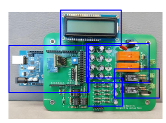
3F 3V Supercap



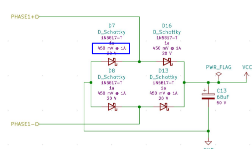
Li-Ion Charging Curve (credit: Battery University)

Electronics Setup

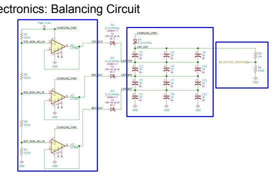




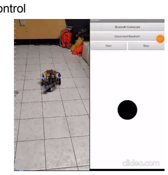
Electronics: Rectifiers



Electronics: Balancing Circuit

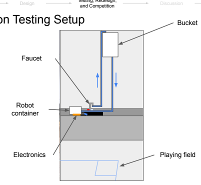


Bluetooth Control



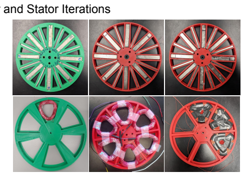
Testing, Redesign, and Competition

Competition Testing Setup

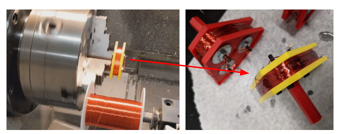


Original Test

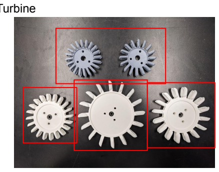
Rotor and Stator Iterations



Stator (cont.)

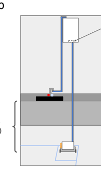


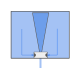
Pelton Turbine



Competition Setup

Result: 48% charging time reduction





The Competition Run

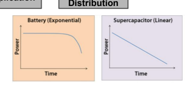
What Went Wrong?

Controllability

Unfortunate Complication

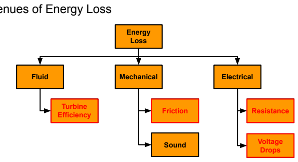
Unequal Weight Distribution

Motor Voltage Threshold

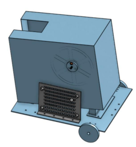


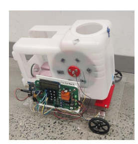
(Credit: Battery University)

Avenues of Energy Loss



Second Iteration of the Cart





Conclusion

Conclusions

Went through **multiple design iterations**

Designed **our own motor controller**

Designed **our own alternator**

Designed **our own PCB**

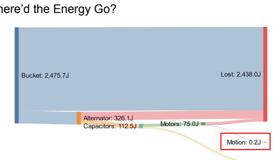
Tripled driving distance

Reduced charging time by **48%**

Tied for **2nd** best design (out of 17 teams)

Thank you for your attention

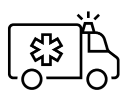
Where’d the Energy Go?



LUNCH

Improving EHR Data Imputation with Distribution-Based Methods

Dan Brody



What is an EHR?

Patient_ID	Date	Prescribed Medications Prednisone	Lab Test Results RBCs	Reported Symptoms Back Pain	...
1	2/3/2019	1	50	1	...
⋮	⋮	⋮	⋮	⋮	

Problem Statement

EHR data is challenging to represent and model due to its high dimensionality, noise, heterogeneity, sparseness, incompleteness, random errors, and systematic biases [1,2,3]. This property of EHR data makes it difficult to use in any deep or machine learning algorithms including algorithms in precision medicine such as endpoint prediction.

What is Endpoint Prediction?

Prescribed Medications				Lab Tests				
Methotrexate	Prednisone	....	Vanomycin	RBC High	RBC Norm	RBC Low	....	WBC Low
1	0		1	1	0	0		1

Methotrexate	Prednisone	....	Vanomycin	RBC High	RBC Norm	RBC Low	....	WBC Low
1	0		0	0	0	1		1

Missingness Types

Probability of missing

- MCAR - unrelated to les and unrelated to the variable itself
- MNAR – dependent ed variables of a variable Y
- MAR – dependent o observed

Prednisone	RBCs
1	1
1	1

Methods:

Median Imputation, GAIN

Median Imputation

Median (

Prednisone
1
0
0
0
1
0
0
0

) =

Fault : 0.5

accuracy depletes as the number of samples in a feature decrease.

GAIN Description [4]

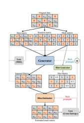


Figure 1. The architecture of GAIN

GAIN Performance [4]

Table 2. Imputation performance in terms of RMSE (Average ± Std of RMSE)

Algorithm	Breast	Spam	Letter	Credit	News
GAIN	**.0546 ± .0006**	**.0513± .0016**	**.1198± .0005**	**.1858 ± .0010**	**.1441 ± .0007**
MICE	.0646 ± .0028	.0699 ± .0010	.1537 ± .0006	.2585 ± .0011	.1763 ± .0007
MissForest	.0608 ± .0013	.0553 ± .0013	.1605 ± .0004	.1976 ± .0015	.1623 ± 0.012
Matrix	.0946 ± .0020	.0542 ± .0006	.1442 ± .0006	.2602 ± .0073	.2282 ± .0005
Auto-encoder	.0697 ± .0018	.0670 ± .0030	.1351 ± .0009	.2388 ± .0005	.1667 ± .0014
EM	.0634 ± .0021	.0712 ± .0012	.1563 ± .0012	.2604 ± .0015	.1912 ± .0011

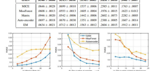


Figure 2. RMSE performance in different settings: (a) Various missing rates, (b) Various number of samples, (c) Various feature dimensions

Table 3. Prediction performance comparison

Algorithm	AUROC (Average ± Std)			
	Breast	Spam	Credit	News
GAIN	**.9930 ± .0073**	**.9529 ± .0023**	**.7527 ± .0031**	**.9711 ± .0027**
MICE	.9914 ± .0034	.9495 ± .0031	.7427 ± .0026	.9451 ± .0037
MissForest	.9860 ± .0112	.9520 ± .0061	.7498 ± .0047	.9597 ± .0043
Matrix	.9897 ± .0042	.8639 ± .0055	.7059 ± .0150	.8578 ± .0125
Auto-encoder	.9916 ± .0059	.9403 ± .0051	.7485 ± .0031	.9321 ± .0058
EM	.9899 ± .0147	.9217 ± .0093	.7390 ± .0079	.8987 ± .0157

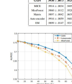


Figure 3. The *AUROC* performance with various missing rates with Credit dataset

Table 4. Congeniality of imputation models

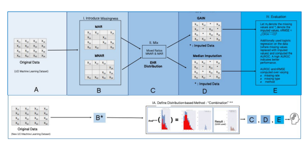
Algorithm	Mean Bias ($\|\|\mathbf{w} - \hat{\mathbf{w}}\|\|_1$)	MSE ($\|\|\mathbf{w} - \hat{\mathbf{w}}\|\|_2$)
GAIN	**0.3163± 0.0887**	**0.5078± 0.1137**
MICE	0.8315 ± 0.2293	0.9467 ± 0.2083
MissForest	0.6730 ± 0.1937	0.7081 ± 0.1625
Matrix	1.5321 ± 0.0017	1.6660 ± 0.0015
Auto-encoder	0.3500 ± 0.1503	0.5608 ±0.1697
EM	0.8418 ± 0.2675	0.9369 ± 0.2296

relationship between feature and label. As can be seen in the table, GAIN achieves significantly lower mean bias and mean square error than other state-of-the-art imputation algorithms (from 8.9% to 79.2% performance improvements).

Methodology

1

2



Results

Note : Missing rate of 0.8 and 50% MNAR and MCAR was assessed. Refer to the paper at [illegible] for more examples and information.

Annotation : The following graph highlights feature-wise nRMSE as well as AUROC for the respective method and dataset where each block of the heatmaps is feature "n" of the dataset

Method	Messidor Heatmap	Messidor AUROC	Credit Card Heatmap	Credit Card AUROC
Combination	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19	0.694	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23	0.603
GAIN	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19	0.695	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23	0.580
Median	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19	0.687	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23	0.69

Conclusions : The "combination" method is consistent across datasets since never gets outperformed by both methods

Future Work

- Using more methods
- Distribution Search API for UCI Machine Learning
- Applying results to EHR data (MIMIC-III) to improve endpoint prediction of myocardial infarction in []

Works Cited

[1] Giraud T, Dhainaut JF, Vaxelaire JF, Joseph T, Journois D, Bleichner G, Sollet JP, Chevret S, Monsallier JF. Iatrogenic complications in adult intensive care units: a prospective two-center study. *Crit Care Med.* 1993;21:40–51. doi: 10.1097/00003246-199301000-00011.

[2]0 Bracco D, Favre JB, Bissonnette B, Wasserfallen JB, Revelly JP, Ravussin P, Chiolero R. Human errors in a multidisciplinary intensive care unit: a 1-year prospective study. *Intensive Care Med.* 2001;27:137–145. doi: 10.1007/s001340000751.

[3] Abramson NS, Wald KS, Grenvik AN, Robinson D, Snyder JV. Adverse occurrences in intensive care units. *JAMA.* 1980;244:1582–1584. doi: 10.1001/jama.1980.03310140040027.

[4]Jinsung Yoon and James Jordon and Mihaela van der Schaar. GAIN: Missing Data Imputation using Generative Adversarial Nets. 2018. GAIN: [1806.02920] GAIN: Missing Data Imputation using Generative Adversarial Nets (arxiv.org)

Applications of Ultra-wideband RTLS in Architecture (AURA)

Theo Song, Dan Park

Applications of Ultra-wideband RTLS in Architecture (AURA)

Donghyun Park
Theo Song

Problems

1. Inefficient architectural design and no systems to analyze it
    a. How is space being used? Bottlenecks?
    b. Maximize human pathing behavior
2. Lack of non-commercial implementation of RTLS at sub-meter accuracy
    a. BLE is well established but inaccurate
3. Ethical dilemmas behind human data collection

Goals: Engineering

- Hardware: Implement small-scale RTLS network system using barebones UWB hardware
- Software: Integrate a web/mobile based application to store RTLS telemetry data, as well as collect user information

Goals: Interdisciplinary

- Collaborate with Architecture School to create visualizations of the aggregated data
- Perform analysis of how space is utilized, how certain groups have similar behaviors, observe any unique patterns in movement etc.

Goals: Social Science

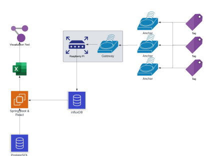
- Discussion of the ethical and moral principles behind tracking technologies applied to humans
- Implicit vs Explicit consent
- Illusion of choice
- Applications of this RTLS systems (beneficial vs detrimental)

Ultra-wideband

+ Great accuracy (10 ~ 20 cm range)
+ Wide Range of 150m
+ Higher data transmission bandwidth
+ Newer generation of localization technology

- Expensive, commercial solutions outside of budget
- Uncommon off-the-shelf parts and commercial implementation

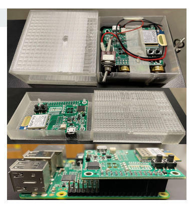
System Stack



Hardware

DWM1001 Development Board

- Emulates both
  - Transmitter/Tag
  - Receiver/Anchor
- One used as gateway to Raspberry Pi using MQTT protocol (used for IoT)



Software: Gateway & Raspberry Pi

- Gateway Built with combination of DWM1001 Board and Raspberry Pi
- Gateway delivers received position information to the MQTT server
- Raspberry Pi using miniconda for Python package management
- Raspberry Pi reads stream of data from MQTT server and stores into influxDB

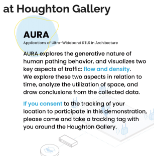
Software: Database

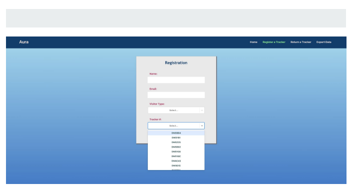
- InfluxDB: time-series database, capable of high flux of insertion and query of time-sequential data
- PostgreSQL: traditional relational DBMS used to store tag and user data
- Both running in AWS

Software: Web Application

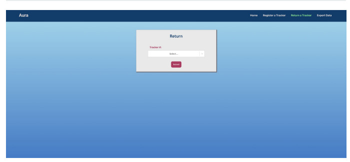
- Backend: Built with Java, Spring Boot, to support following APIs
  - Assigning unoccupied tags to incoming users
  - Detaching tags for leaving users
  - Export Position Data in a form of Excel file
- Frontend: Built with React for easy access of functionalities

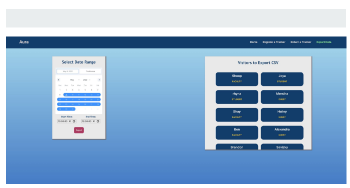
Demo Run at Houghton Gallery

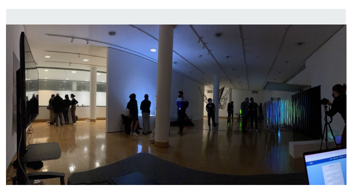


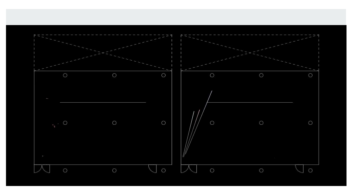


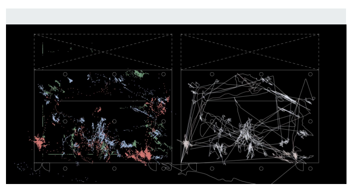
Aura
Home
Register a Tracker
Return a Tracker
Export Data
Registration
Name:
Email:
Visitor Type:
Select...
Tracker #:
Select...


Aura
Home
Register a Tracker
Return a Tracker
Export Data
Return
Tracker #:
Select...
Submit


Aura
Home
Register a Tracker
Return a Tracker
Export Data
Select Date Range
Start Time:
End Time:
Export
Visitors to Export CSV
Shoop
FACULTY
Joya
STUDENT
rhyna
STUDENT
Mersiha
GUEST
Shay
FACULTY
Hailey
GUEST
Ben
FACULTY
Alexandra
GUEST
Brandon
Savizky







Social Science

- Benevolent Uses
  - Improving floor plans of frequently congested areas
  - Better public infrastructure
    - public fountains, bathrooms, info centers etc.
- Malevolent Uses
  - Collected information could be sold as data
  - Invasion of privacy in exchange for access to service
- Better awareness from both users and providers are required

Future Work

- Software
    - Scaling the system to support massive numbers of users, and process the increased data
    - More personal information to perform more rigorous data analysis
        - Ethnic patterns
        - Age patterns
        - Gender patterns
- Interdisciplinary
    - Create better visualization pipeline with more robust tools to remove computational bottleneck
    - Come up with a systemic approach to redesign floor plans to be more effective through means such as machine learning

Acknowledgements

- Professor Keene
- School of Architecture
    - Kyungmin Park
    - Tianyang Sun

CARBON CURATOR

By: Danny Hong and Arthur Skok

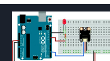
Problem Statement

Scientific advancements have separated many people from the environment they live in.

Many people perform tasks that have harmful effects on the environment without realizing that there are more eco friendly alternative behaviors to do.

We propose a portable device that can help users become more aware of how their actions have

What is it composed of?



The project is composed of a Hardware and Software Component

- **Hardware Component**
  - Consists of an Arduino which is hooked up to the Temperature/Humidity, and GPS sensors.
  - Also has a battery pack connected to get power.
  - Has a Bluetooth component in order to have internet connection between the Arduino and the code.
- **Software Component**
  - Consists of the user interface experience between the user and the device.
  - Implements storage for user inputted data combined with sensory data

Current Iteration: Hardware Specs

ARDUINO SPECS	TEMPERATURE SENSOR SPECS	GPS SENSOR SPECS
- 256 KB of flash memory 8 KB used by bootloader - 16 MHz clock speed - 54 digital input/output pins - 5 V operating voltage - 37 grams of weight, 53 mm by 100 mm size (for portability)	- 3.3 to 6 V DC power supply - 0 - 100% relative humidity operating range - 0.1% relative humidity resolution (precision) - +/- 2% relative humidity accuracy - -40 to 80° C operating range - 0.1° C temperature precision - +/- 0.5° C temperature accuracy	- 3.6 V power supply - 50 db antenna gain - 9600 baud rate - 2.5 meter precision

Current Iteration: Software

Communication Between Devices

- Data is encoded into packets on the arduino side of the code
  - Broadcasted via the usage of characteristics on the HM-10 BLE device
  - Packets decoded on application side

Data Utilization

- After decoding, data is stored locally and treated differently based on packet type:
  - Temperature data is parsed into temperature and humidity readings
  - GPS data is parsed into longitude and latitude
- Distance accumulates and resets, real outside temperature fetched from openweather API key and compared to readings
- Depending on what values are calculated, states are set throughout the application for rendering and alerts

Current Iteration: Software (Continued)

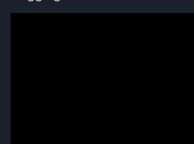
Communicating to the User

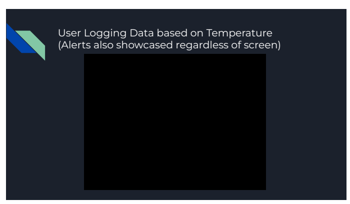
- If specific thresholds are met, a notification via alert is sent to the user to tell them that a question is waiting for them at the prompt screen
- Furthermore, the navigation button rerenders as an exclamation point until the user answers the program in affirmative or negative to signify pending questions
    - Depending on the user response, the application seeks further clarification for what the user is doing

Storing Responses

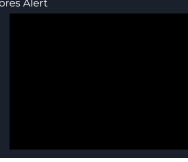
- Certain answers are then stored locally
    - However this data needs to stay in memory even as the application goes out of scope, turned off, or anything shy of an uninstall
- We decided to implement SQLite for our application for persistent data storage in this case
- The data being stored in the SQLite database is rendered to the user in the prompt screen so they may see their previous results

User Logging Data based on Travel

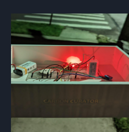



User Logging Data based on Temperature
(Alerts also showcased regardless of screen)

Pending Prompt Changes Render Status if User Ignores Alert



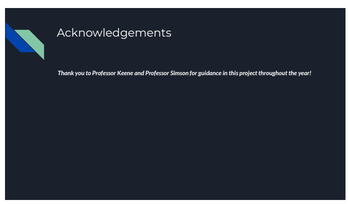
...And it's portable!



...sort of

What is next?

- **Additional user prompts:**
  - What are you doing?
    - Purchases, working
- **Recommendations:**
  - Alternative transit options:
    - Fetch nearby transit features from google maps API using stored GPS data at user end travel entries
  - Alternative purchasing options:
    - Goods can be hard coded to give an environmental grade
  - Reducing excessive heating or cooling
- **Hardware:**
  - Making our design more compact and portable
    - Further refining our physical carriage to take up less space, perhaps by mounting components facing away from each other



Acknowledgements
Thank you to Professor Keene and Professor Simson for guidance in this project throughout the year!

A Hybrid Approach for Image Vectorization for Semi-Geometric Images

by
Jonathan Lam, Derek Lee, Victor Zhang

Prof. Sam Keene

May 13, 2022

Problem statement

- Converting a raster (pixel-based) image to vector (shape-based) image
- Develop hybrid method that combines benefits of previous methods

Edge tracing

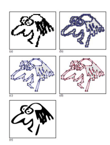
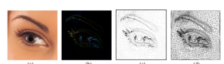


Figure: Illustration of the Potrace [1] vectorization process

Blue-noise sampling



(a) (b) (c) (d)

Figure: Illustration of the BNS vectorization process [2]

Hybrid approach

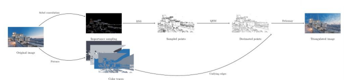
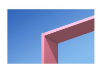
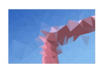


Figure: Architecture diagram

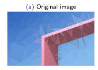
Results



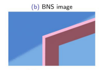
(a) Original image



(b) BNS image



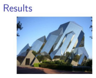
(c) Hybrid image



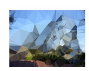
(d) Potrace image

Figure: Set of images for experiment 3

Results

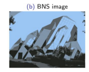


(a) Original image



(b) BNS image

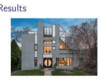
(c) Hybrid image



(d) Potrace image

Figure: Set of images for experiment 4

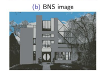
Results



(a) Original image

(b) BNS image

(c) Hybrid image



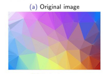
(d) Potrace image

Figure: Set of images for experiment 5

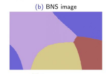
Results

(a) Original image

(b) BNS image



(c) Hybrid image



(d) Potrace image

Figure: Set of images for experiment 8

Future work

- Alternative methods to strengthen edges
- Curve simplification
- Machine learning preprocessing
- Mathematical model of pipeline
- Improved evaluation metrics

Conclusions

- Implemented framework for vectorizing images
- Based on blue-noise sampling and Potrace
- Larger file size in exchange for better performance on accuracy (MSE)

References

[1] Peter Selinger. “Potrace: a polygon-based tracing algorithm”. In: *Potrace (online), http://potrace.sourceforge.net/potrace.pdf (2009-07-01)* 2 (2003).

[2] Jiaojiao Zhao, Jie Feng, and Bingfeng Zhou. “Image vectorization using blue-noise sampling”. In: *Imaging and Printing in a Web 2.0 World IV*. Vol. 8664. International Society for Optics and Photonics. 2013, 86640H.