Control, Modeling and Computing

David C. Wyld,
Dhinaharan Nagamalai (Eds)

Computer Science & Information Technology

- 8th International Conference on Control, Modeling and Computing (CMC 2022), March 19~20, 2022, Vienna, Austria
- 8th International Conference on Software Engineering (SOFT 2022)
- 3rd International Conference on Machine Learning &Trends (MLT 2022)
- 11th International Conference on Soft Computing, Artificial Intelligence and Applications (SAI 2022)

Published By

AIRCC Publishing Corporation

Volume Editors

David C. Wyld,
Southeastern Louisiana University, USA
E-mail: David.Wyld@selu.edu

Dhinaharan Nagamalai (Eds),
Wireilla Net Solutions, Australia
E-mail: dhinthia@yahoo.com

ISSN: 2231 - 5403
ISBN: 978-1-925953-64-0
DOI: 10.5121/csit.2022.120501- 10.5121/csit.2022.120505

Typesetting: Camera-ready by author, data conversion by NnN Net Solutions Private Ltd., Chennai, India

Preface

8th International Conference on Control, Modeling and Computing (CMC 2022), March 19~20, 2022, Vienna, Austria, 8th International Conference on Software Engineering (SOFT 2022), 3rd International Conference on Machine Learning &Trends (MLT 2022) and 11th International Conference on Soft Computing, Artificial Intelligence and Applications (SAI 2022) was collocated with 8th International Conference on Control, Modeling and Computing (CMC 2022). The conferences attracted many local and international delegates, presenting a balanced mixture of intellect from the East and from the West.

The goal of this conference series is to bring together researchers and practitioners from academia and industry to focus on understanding computer science and information technology and to establish new collaborations in these areas. Authors are invited to contribute to the conference by submitting articles that illustrate research results, projects, survey work and industrial experiences describing significant advances in all areas of computer science and information technology.

The CMC 2022, SOFT 2022, MLT 2022 and SAI 2022 Committees rigorously invited submissions for many months from researchers, scientists, engineers, students and practitioners related to the relevant themes and tracks of the workshop. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer review process which comprised expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The papers were then reviewed based on their contributions, technical content, originality and clarity. The entire process, which includes the submission, review and acceptance processes, was done electronically.

In closing, CMC 2022, SOFT 2022, MLT 2022 and SAI 2022 brought together researchers, scientists, engineers, students and practitioners to exchange and share their experiences, new ideas and research results in all aspects of the main workshop themes and tracks, and to discuss the practical challenges encountered and the solutions adopted. The book is organized as a collection of papers from the CMC 2022, SOFT 2022, MLT 2022 and SAI 2022.

We would like to thank the General and Program Chairs, organization staff, the members of the Technical Program Committees and external reviewers for their excellent and tireless work. We sincerely wish that all attendees benefited scientifically from the conference and wish them every success in their research. It is the humble wish of the conference organizers that the professional dialogue among the researchers, scientists, engineers, students and educators continues beyond the event and that the friendships and collaborations forged will linger and prosper for many years to come.

David C. Wyld,
Dhinaharan Nagamalai (Eds)

General Chair

Organization

David C. Wyld, Southeastern Louisiana University, USA
Dhinaharan Nagamalai (Eds) Wireilla Net Solutions, Australia

Program Committee Members

Abdel-Badeeh M. Salem, Ain Shams University, Egypt
Abderrahmane Ez-Zahout, Mohammed V University, Morocco
Abdessamad Belangour, University Hassan II Casablanca, Morocco
Adrian Olaru, University Politehnica of Bucharest, Romania
Afaq Ahmad, Sultan Qaboos University, Muscat, Oman
Ahmed Moumena, Hassiba benbouali university of chlef, Algeria
Akhil Gupta, Lovely Professional University, India
Ali Fadhil Naser, Al-Furat Al-Awsat Technical University, Iraq
Alireza Valipour Baboli, University Technical and Vocational, Babol, Iran
Allel Hadjali, LIAS/ENSMA, France
Amir H Gandomi, University of Technology, Australia
Andy Rachman, Institut Teknologi Adhi Tama Surabaya, Indonesia
Anouar Abtoy, Abdelmalek Essaadi University, Morocco
António Abreu, ISEL – Polytechnic Institute of Lisbon, Portugal
Cheng Siong Chin, Newcastle University, Singapore
D Lakshmi Padmaja, Anurag University, India
Daniel Hunyadi, "Lucian Blaga" University of Sibiu, Romania
Dário Ferreira, University of Beira Interior, Portugal
Dariusz Jacek Jakobczak, Koszalin University of Technology, Poland
Debjani Chakraborty, Indian Institute of Technology, India
Dimiter Velev, University of National and World Economy, Bulgaria
Dimitris Kanellopoulos, University of Patras, Greece
Djenouhat, University of Algiers 3, Algeria
El Murabet Amina, Abdelmalek Essaadi University, Morocco
Elżbieta Macioszek, Silesian University of Technology, Poland
Eugénia Moreira Bernardino, Polytechnic Institute of Leiria, Portugal
Felix J. Garcia Clemente, University of Murcia, Spain
Fernando Zacarias Flores, Universidad Autonoma de Puebla, Mexico
Francesco Zirilli, Sapienza Universita Roma , Italy
Gourav Bansal, RCAM, United States
Grigorios N. Beligiannis, University of Patras, Greece
Grzegorz Sierpiński, Silesian University of Technology, Poland
Hamid Ali Abed AL-Asadi, Iraq University college, Iraq
Hamidreza Rokhsati, Sapienza University of Rome, Italy
Hedayat Omidvar, National Iranian Gas Company, Iran
Ilango Velchamy, CMR Institute of Technology, India
Israa Shaker Tawfic, Ministry of Migration and Displaced, Iraq
Javier Gozalvez, Universidad Miguel Hernandez de Elche, Spain
Jesuk Ko, Universidad Mayor de San Andres, Bolivia
Kamel Hussein Rahouma, Nahda University,Beni Suef, Egypt
Kazuyuki Matsumoto, Tokushima University, Japan
Khaled osama Elzoghaly, Alexandria University, Egypt
Loc Nguyen, Loc Nguyen's Academic Network, Vietnam
Luisa Maria Arvide Cambra, University of Almeria, Spain

Technically Sponsored by

Computer Science & Information Technology Community (CSITC)

Artificial Intelligence Community (AIC)

Soft Computing Community (SCC)

Digital Signal & Image Processing Community (DSIPC)

8th International Conference on Control, Modeling and Computing (CMC 2022)

8th International Conference on Software Engineering (SOFT 2022)

3rd International Conference on Machine Learning &Trends (MLT 2022)

11th International Conference on Soft Computing, Artificial Intelligence and Applications (SAI 2022)

STATE DRIFT AND GAIT PLAN IN FEEDBACK LINEARIZATION CONTROL OF A TILT VEHICLE

Zhe Shen and Takeshi Tsuchiya

Department of Aeronautics and Astronautics,
The University of Tokyo, Tokyo, Japan

ABSTRACT

To stabilize a conventional quadrotor, simplified equivalent vehicles (e.g., autonomous car) are developed to test the designed controller. Based on that, various controllers based on feedback linearization have been developed. With the recently developed concept of tilt-rotor, there lacks the simplified/equivalent model, however. Indeed, the tilt structure is relatively unusual in vehicles. In this research, we put forward a unique fictional vehicle with tilt structure, which is to help evaluate the property of the tilt-structure-aimed controllers. One phenomenon (state drift) in controlling an over-actuated tilt structure by feedback linearization is presented subsequently. State drift can be easily neglected and is not paid attention to in the current researches in tilt-rotor controllers' design so far. We report this phenomenon and provide a potential approach to avoid this behavior.

KEYWORDS

Feedback Linearization, State Drift, Over-actuated System, Gait Plan, Stability.

1. INTRODUCTION

Various controllers are analyzed to control the conventional quadrotors. Some of them are linear. [5] compares two linear controllers in stabilizing the attitude of a quadrotor model. Others are nonlinear. [6, 7] apply geometric controller, which delicately defines the attitude error and provides the stability proof. [8] puts forward a sliding mode controller; it guarantees the stability using Lyapunov criteria. Backstepping controller with cascade structure can be found in [9].

Both linear and nonlinear controllers have their advantages and disadvantages. These pros and cons include the controlling time, stability region, dynamic bias, robustness, calculating cost, etc. Different from solely relying on linear or nonlinear controllers, feedback linearization [2,3] transfers a nonlinear system to a linear system. The linear controller can be subsequently applied to the generated linear system.

Unfortunately, the quadrotor is an under-actuated system where the number of inputs is less than the degrees of freedom. This nature prohibits us from building a full-state controller; the number of the states to be controlled directly is 4 at most, equal to the number of the inputs. Several different selections on these states can be found in [2, 10, 11]. Another problem hindering us from utilizing feedback linearization is the invertibility of the decouple matrix [2, 12]. Some selections of the output combination introduce the singular decouple matrix. [12] details this problem and provides the potential approach to avoid hitting the singular zone.

David C. Wyld et al. (Eds): CMC, SOFT, MLT, SAI - 2022
pp. 01-17, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.120501

The recent popular novel quadrotor (tilt-rotor) augments the number of inputs to 8 [4, 13, 14]; the quadrotor not only changes the magnitude of each thrust but also the direction. With this designation, the system becomes over-actuated. Subsequently, several researches focus on controlling this over-actuated system by feedback linearization [4], [15]-[17].

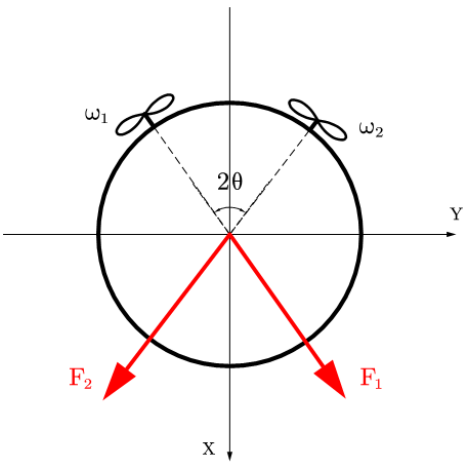
Controlling an over-actuated system by feedback linearization provides the possibility of full-state control. The input, however, might change even if the state has stabilized. This phenomenon is not paid attention to yet.

In this paper, we address the feedback-linearization-based controllers in a simplified tilt vehicle. Controlling a simplified vehicle is a common approach before applying this control method to the complicated system (e.g., UAV) [1]. We present the result with the state drift phenomenon. To avoid this unwanted drift, we decrease the number of the real inputs by gait plan, which is a gait schedule technique in quadruped robots [18].

This paper is structured as follows: Section 2 introduces the dynamics of the fictional vehicle. The feedback-linearization-based controller is designed in Section 3. Section 4 simulates the controller designed and presents the state drift phenomenon. Section 5 proposes gait plan and puts it into the controller design. Section 6 concludes and makes the discussion of the result.

2. DYNAMICS OF THE TILT VEHICLE

In this section, we introduce a two-dimensional fictional tilt vehicle (Figure 1). The motion of the vehicle is restricted to a level plane (two-dimensional movement). Two propellers are fixed on the vehicle with the angle 2θ between. The value of 2θ is $\frac{\pi}{3}$.



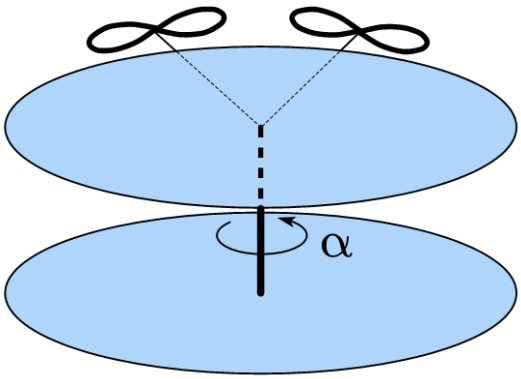


Figure. 1. Two-dimensional Tilt Vehicle Figure. 2. Tilt Structure

Each propeller can generate the thrust by rotating with angular velocity, ω_1 and ω_2. The trust generated follows Equation (1).

$$F_i = K_{F_i} \cdot \omega_i^2,\ i = 1, 2\ \ (\omega_i \geq 0) \tag{1}$$

where K_{F_i} is the thrust coefficient. Its value is $K_{F_i} = 0.001$. This proportional relationship in adopted in the dynamics of UAV control [1-3]. The thrust is always nonnegative. Also, the angular velocity is always nonnegative.

What makes this vehicle special is the tilt structure, as plotted in Figure 2. The propellers mentioned are fixed on the top disc. While the bottom disc contacts the $x-y$ plane directly. The frictional force from the $x-y$ plane is ignored. This structure provides the possibility of changing the directions of the thrusts by tilting the top disc (Figure 3—4) with an angle, α.

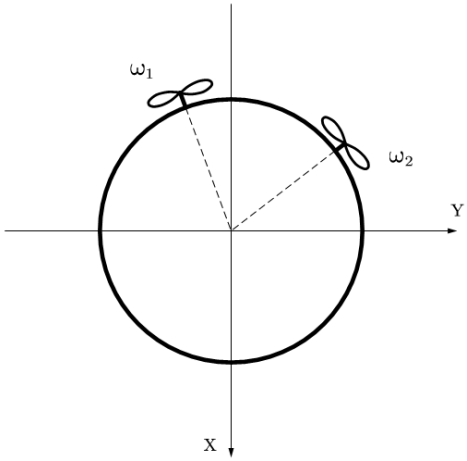


Fig. 3. Tilt to right example

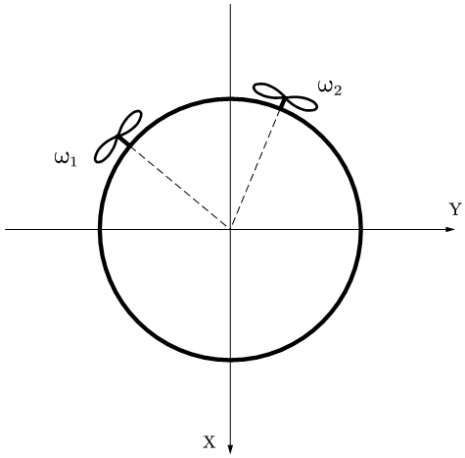


Fig. 4. Tilt to left example

Based on Conservation of Angular Momentum, the relationship between the yaw, ψ, and the tilt angle, α, is written in Equation (2).

$$\psi = -\frac{I_T}{I_B} \cdot \alpha \tag{2}$$

where I_T and I_B are the rotational inertias of the top disc and bottom disc, respectively. Their values are $I_T = 0.001$, $I_B = 0.002$.

Based on the Newtown's law, we write the position in Equation (3).

$$\begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix} = \frac{1}{m} \cdot J_{\psi+\alpha} \cdot J_{\theta} \cdot \begin{bmatrix} {\omega_1}^2 \\ {\omega_2}^2 \end{bmatrix} \tag{3}$$

where m is the mass of the entire vehicle. $J_{\psi+\alpha}$ and J_{θ} are defined in Equation (4)—(5).

$$J_{\psi+\alpha} = \begin{bmatrix} \cos(\psi+\alpha) & -\sin(\psi+\alpha) \\ \sin(\psi+\alpha) & \cos(\psi+\alpha) \end{bmatrix} \tag{4}$$

$$J_{\theta} = \begin{bmatrix} \cos(\theta) & 0 \\ 0 & \sin(\theta) \end{bmatrix} \cdot \begin{bmatrix} K_{F_1} & K_{F_2} \\ K_{F_1} & -K_{F_2} \end{bmatrix} \tag{5}$$

It is worth mentioning that both $J_{\psi+\alpha}$ and J_θ have full rank. Thus, $J_{\psi+\alpha} \cdot J_\theta$ is invertible.

$$|J_{\psi+\alpha} \cdot J_\theta| \neq 0 \tag{6}$$

So far, we have derived the dynamics of this vehicle in Equation (2)—(3). This is a MIMO system. There are three inputs, ω_1, ω_2, and α. While there are two outputs, x, and y. So, it is an over-actuated system.

3. CONTROLLER DESIGN (FEEDBACK LINEARIZATION)

Remarks to this MIMO system: Since it is an over-actuated system, we may apply feedback linearization for this system. The other idea is to discard some inputs to let the number of the outputs equal to the number of the inputs. Although Feedback Linearization is widely used the tilt-rotor control, one potential adverse effect (state drift) is always neglected by researchers.

This section develops the feedback linearization controllers based on this over-actuated system model. The result and the state drift phenomenon are given in the next section.

3.1. Third Derivative Feedback Linearization

Since the input α is tangled in Equation (3), we calculate the higher derivative (third order here) of the position in Equation (7).

$$\begin{bmatrix} \dddot{x} \\ \dddot{y} \end{bmatrix} = \frac{1}{m} \cdot \dot{J}_{\psi+\alpha} \cdot J_\theta \cdot \begin{bmatrix} {\omega_1}^2 \\ {\omega_2}^2 \end{bmatrix} + \frac{1}{m} \cdot J_{\psi+\alpha} \cdot J_\theta \cdot \begin{bmatrix} 2\omega_1 \cdot \dot{\omega}_1 \\ 2\omega_2 \cdot \dot{\omega}_2 \end{bmatrix} \tag{7}$$

where $\dot{J}_{\psi+\alpha}$ is calculated in Equation (8).

$$\dot{J}_{\psi+\alpha} = \begin{bmatrix} -\sin(\psi+\alpha) & -\cos(\psi+\alpha) \\ \cos(\psi+\alpha) & -\sin(\psi+\alpha) \end{bmatrix} \cdot \left(\dot{\psi}+\dot{\alpha}\right) \tag{8}$$

Substituting Equation (2) into Equation (7) yields Equation (9).

$$\begin{bmatrix} \dddot{x} \\ \dddot{y} \end{bmatrix} = [{\Delta_1}^{2\times 2} | {\Delta_2}^{2\times 1}] \cdot \begin{bmatrix} \dot{\omega}_1 \\ \dot{\omega}_2 \\ \dot{\alpha} \end{bmatrix} \tag{9}$$

where ${\Delta_1}^{2\times 2}$ and ${\Delta_2}^{2\times 1}$ are defined in Equation (10)—(11).

$${\Delta_1}^{2\times 2} = \frac{1}{m} \cdot J_{\psi+\alpha} \cdot J_\theta \cdot \begin{bmatrix} 2\omega_1 & 0 \\ 0 & 2\omega_2 \end{bmatrix} \tag{10}$$

$${\Delta_2}^{2\times 1} = \frac{1}{m} \cdot \begin{bmatrix} -\sin(\psi+\alpha) & -\cos(\psi+\alpha) \\ \cos(\psi+\alpha) & -\sin(\psi+\alpha) \end{bmatrix} \cdot J_\theta \cdot \left(1 - \frac{I_T}{I_B}\right) \cdot \begin{bmatrix} {\omega_1}^2 \\ {\omega_2}^2 \end{bmatrix} \tag{11}$$

Since ${\Delta_1}^{2\times 2}$ is invertible if and only if $\omega_1 \neq 0$ and $\omega_2 \neq 0$, $[{\Delta_1}^{2\times 2}|{\Delta_2}^{2\times 1}]$ has full row rank if and only if $\omega_1 \neq 0$ and $\omega_2 \neq 0$. Thus, the pseudo-inverse of $[{\Delta_1}^{2\times 2}|{\Delta_2}^{2\times 1}]$ has explicit form, Equation (12).

$$pinv(\Delta) = \Delta^T \cdot (\Delta \cdot \Delta^T)^{-1} \tag{12}$$

where $\Delta = [{\Delta_1}^{2\times 2}|{\Delta_2}^{2\times 1}]$.

Separating the inputs in Equation (9), we receive Equation (13).

$$\begin{bmatrix} \dot{\omega}_1 \\ \dot{\omega}_2 \\ \dot{\alpha} \end{bmatrix} = pinv([{\Delta_1}^{2\times 2}|{\Delta_2}^{2\times 1}]) \cdot \begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix} \tag{13}$$

Based on Equation (13), the controller is designed in Equation (14).

$$\begin{bmatrix} \dot{\omega}_1 \\ \dot{\omega}_2 \\ \dot{\alpha} \end{bmatrix} = pinv([{\Delta_1}^{2\times 2}|{\Delta_2}^{2\times 1}]) \cdot \begin{bmatrix} \ddot{x}_r + k_{x_1} \cdot (\ddot{x}_r - \ddot{x}) + k_{x_2} \cdot (\dot{x}_r - \dot{x}) + k_{x_3} \cdot (x_r - x) \\ \ddot{y}_r + k_{y_1} \cdot (\ddot{y}_r - \ddot{y}) + k_{y_2} \cdot (\dot{y}_r - \dot{y}) + k_{y_3} \cdot (y_r - y) \end{bmatrix} \tag{14}$$

where x_r and y_r are the position reference. k_{x_i} and k_{y_i} $(i = 1, 2, 3)$ are the coefficients for the controller.

4. SIMULATION AND STATE DRIFT PHENOMENON

This section provides the simulation results of the control rules in Section 3. An interesting phenomenon, which we call state drift, in the result is paid attention to.

The reference is a circular trajectory starting from $(0, 0)$ in Figure 5. The radius of the reference is 10. The velocity of the reference is 10. The initial position of the vehicle is $(0, 0)$. The initial velocity, is also zero. The initial angular velocities are $\omega_1 = 200$ and $\omega_2 = 200$.

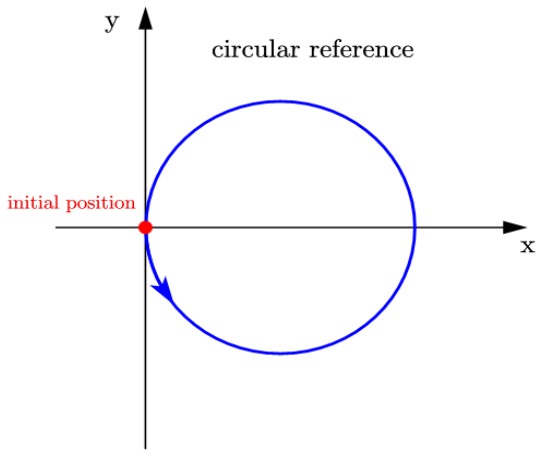


Fig. 5. The circular reference.

The solver is a fixed-step ODE3 in Simulink, MATLAB. The sampling time is 0.01 second.

4.1. Results for Third Derivative Feedback Linearization

The Simulink block diagram is plotted in Figure 6. The Dynamics part, Reference part, and Controller part are also shown.

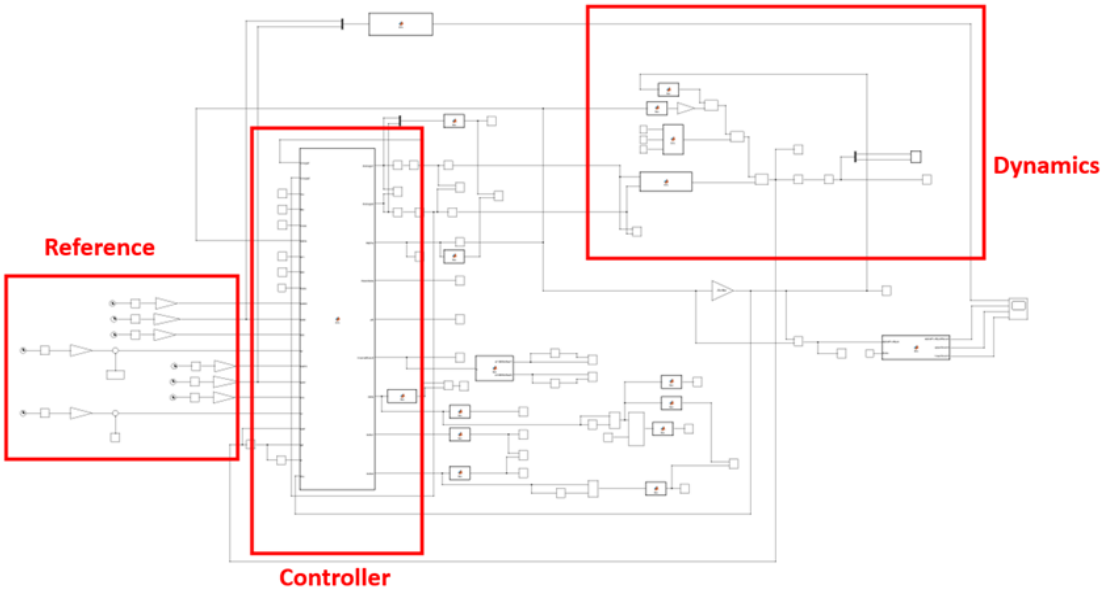


Fig. 6. Simulink block in simulation

Set the simulation time 10 seconds. The results are illustrated in Figure 7—10.

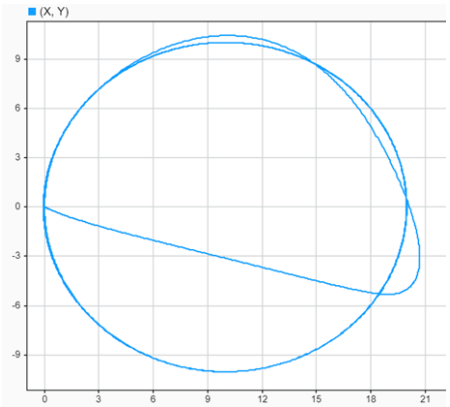


Fig. 7. The trajectory.

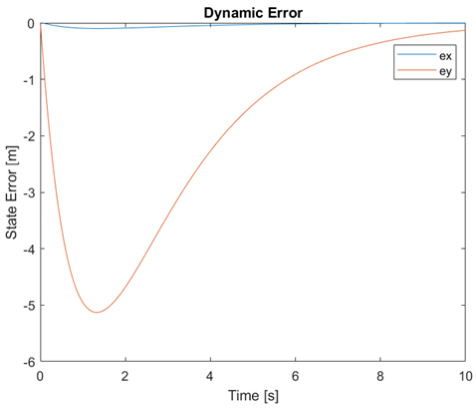


Fig. 8. The dynamic state error.

In Figure 7, the vehicle moves on a biased circular trajectory. The dynamic state error can be checked in Figure 8. Both the dynamic state errors in x and y directions decrease to zero.

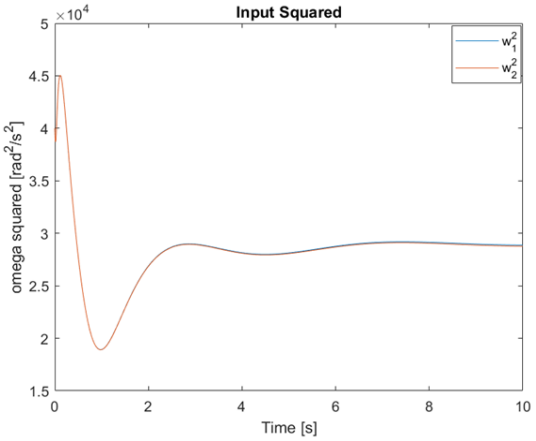


Fig. 9. Input squared.

Figure 9 shows the square of the input signals. Notice that two inputs almost overlap each other in Figure 9, especially at the beginning. The input signals seem to diverge/separate in the later time.

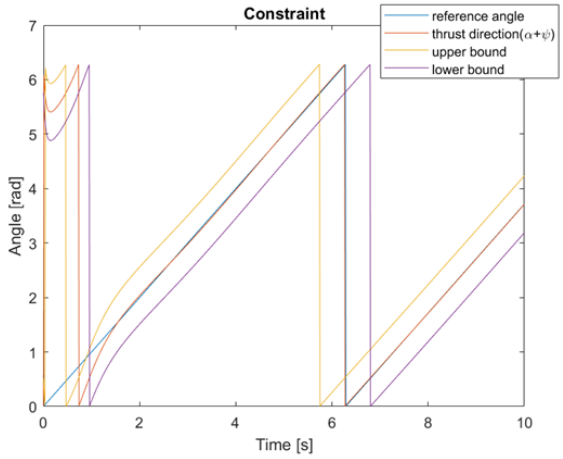


Fig. 10. The direction of the thrust.

Figure 10 plots the information related to the direction of the thrust. The direction of the middle of two thrusts can be calculated in Equation (15).

$$\alpha + \psi \tag{15}$$

The orange line in Figure 10 represents the direction of the middle of two thrusts. The yellow line (upper bound) and the purple line (lower bound) represent the admissible direction bound of the vehicle. The potential direction for generating the acceleration is between the lower bound and the upper bound. The upper bound and lower bound are computed in Equation (16)—(17), respectively.

$$\alpha + \psi + \frac{\pi}{6} \tag{16}$$

$$\alpha + \psi - \frac{\pi}{6} \tag{17}$$

The blue line in Figure 10 is the direction pointing to the center of the circle (Figure 11). We can see that the direction of the middle of two thrusts of the vehicle (orange line) seems to stabilize at the direction pointing to the center of the circle (Figure 11) when time approaches 10 seconds.

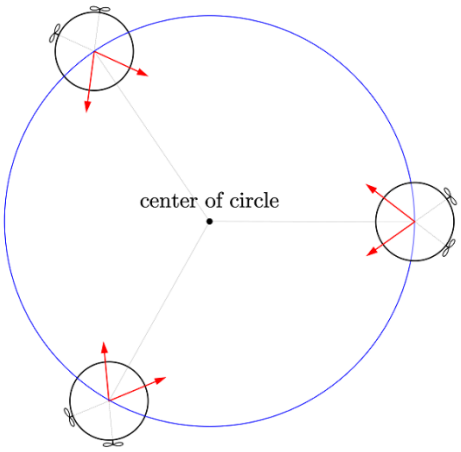


Fig. 11. The direction pointing to the center of the circle

Extending the simulation time to 2000 seconds, we receive the results in Figure 12—15.

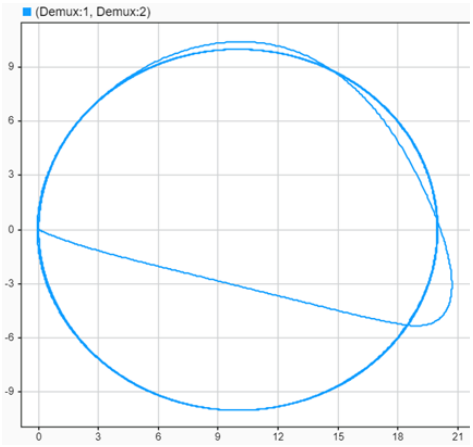


Fig. 12. The trajectory.

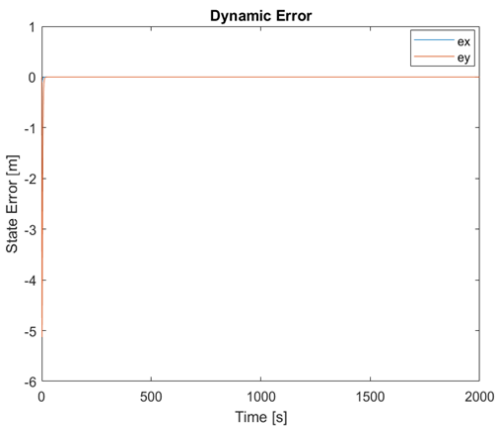


Fig. 13. The dynamic state error.

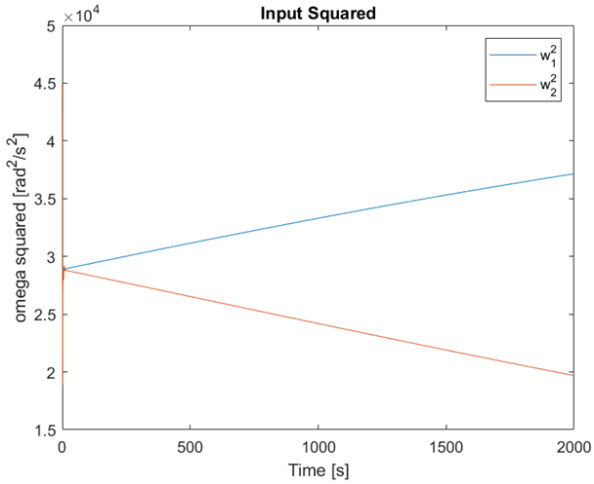


Fig. 14. Input squared.

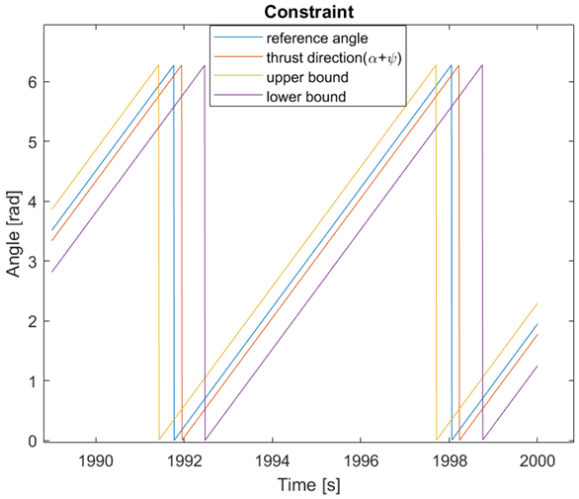


Fig. 15. The direction of the thrust.

The trajectory follows the circle in high precision (Figure 12,13) after being controlled for sufficient long time. The dynamic state error remains close to zero, indicating the stability of the system. Figure 14 shows the squares of the two angular velocities. The inputs diverge in two opposite directions.

Figure 15 shows the acceleration constraint around 1990—2000 seconds. The blue line in Figure 15 is the direction pointing to the center of the circle (Figure 11). The orange line is the direction of the middle of two thrusts. It dramatically biases from the direction pointing to the center of the circle (blue line).

4.2. State Drift

The interesting phenomenon in the result is that the input changes dramatically after the vehicle is stable already. We define this phenomenon ‘State Drift’. State Drift happens only in the over-actuated systems. The direct reason causing State Drift refers to Equation (12), (14). While deducing it mathematically in a systematic way can be hard or even impossible.

While the underlying reason causing State Drift in this vehicle can be obviously explained.

The centripetal force pointing to the center of the circle is required to follow the reference in Figure 15. While generating a specific centripetal force, the combination of the magnitudes of the two thrusts (F_1, F_2) is not unique (Figure 16—18). State Drift can be avoided if the number of the inputs is no more than the number of the output.

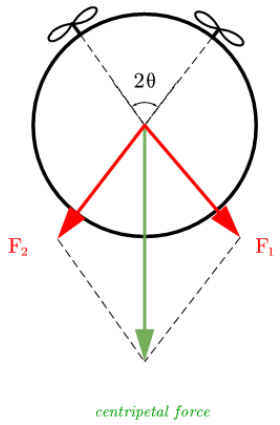


Fig. 16. Symmetrical.

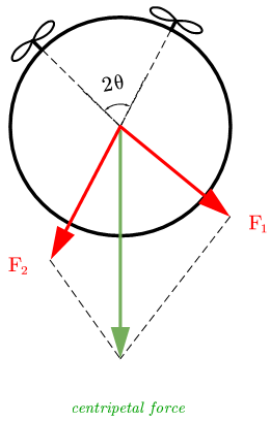


Fig. 17. Symmetrical.

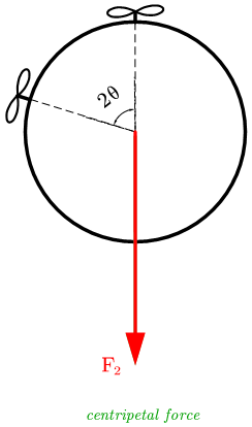


Fig. 18. Degrade to one thrust.

5. GAIT PLAN

The approach to avoid the State Drift is to decrease the number of inputs to be equal to the number of the outputs. In this research, we plan the direction of the middle of two thrusts in advance and discard the tilt angle, α, from the input. The only two inputs then are the angular velocities of the two propellers, ω_1 and ω_2.

α is defined in Equation (18) such that the direction of the middle of two thrusts $(\psi + \alpha)$ always points to the center of the circular reference in Figure 11 during the entire movement.

$$\alpha = \frac{1}{1 - \frac{I_T}{T_B}} \cdot t \tag{18}$$

Having decided α in Equation (18), only two inputs, ω_1 and ω_2, are left to be defined.

Based on the dynamics in Equation (3), the dynamic inversion is completed in Equation (19).

$$\begin{bmatrix} {\omega_1}^2 \\ {\omega_2}^2 \end{bmatrix} = m \cdot \left(J_{\psi+\alpha} \cdot J_\theta\right)^{-1} \cdot \begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix} \tag{19}$$

It is worth mentioning that Equation (19) can be received for the reason that the term $J_{\psi+\alpha} \cdot J_\theta$ is invertible.

Further, we develop the PD controllers in Equation (20).

$$\begin{bmatrix} {\omega_1}^2 \\ {\omega_2}^2 \end{bmatrix} = m \cdot \left(J_{\psi+\alpha} \cdot J_\theta\right)^{-1} \cdot \begin{bmatrix} \ddot{x}_r + k_{x_1} \cdot (\dot{x}_r - \dot{x}) + k_{x_2} \cdot (x_r - x) \\ \ddot{y}_r + k_{y_1} \cdot (\dot{y}_r - \dot{y}) + k_{y_2} \cdot (y_r - y) \end{bmatrix} \quad (20)$$

where x_r and y_r are the position reference. k_{x_i} and k_{y_i} $(i = 1, 2)$ are the coefficients for the controller.

The same reference in Figure 5 is used for the tracking experiment to test the controller in Equation (20). The Simulink block diagram for this experiment is plotted in Figure 19. The dynamics, controller, and the reference parts are marked.

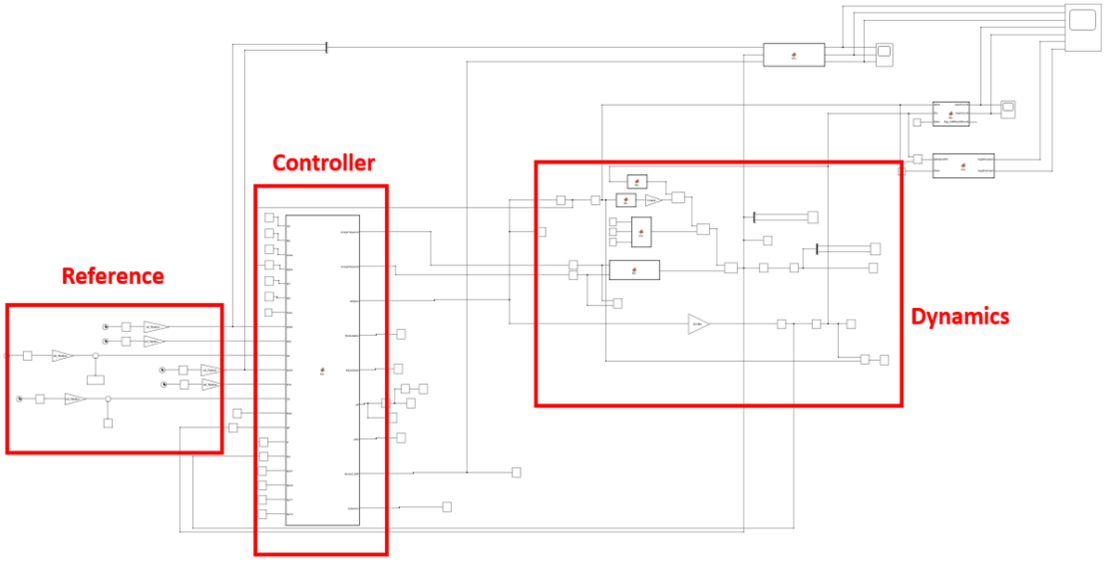


Fig. 19. Simulink block diagram

The simulation time is 10 seconds. The sampling time is 0.01 second (ODE3 fixed-step solver). The result is plotted in Figures 20—23.

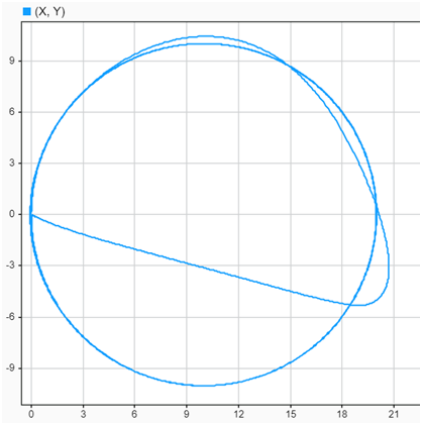


Fig. 20. The trajectory.

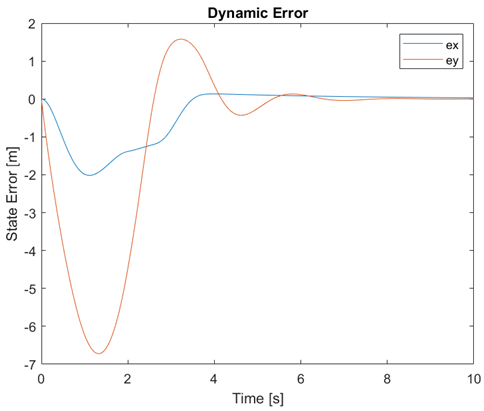


Fig. 21. The dynamic state error.

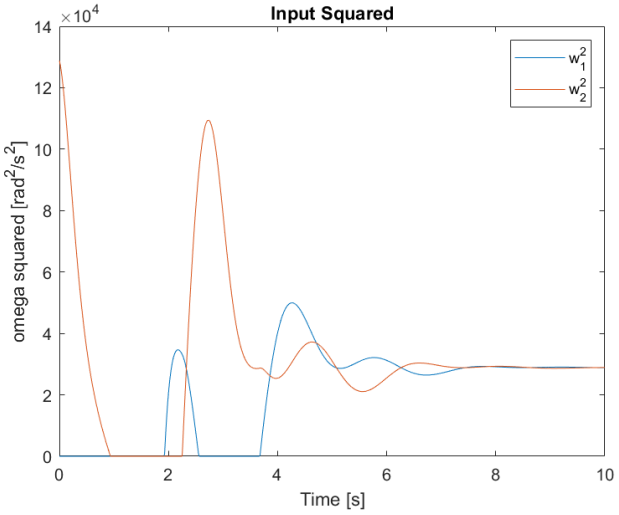


Fig. 22. Input squared.

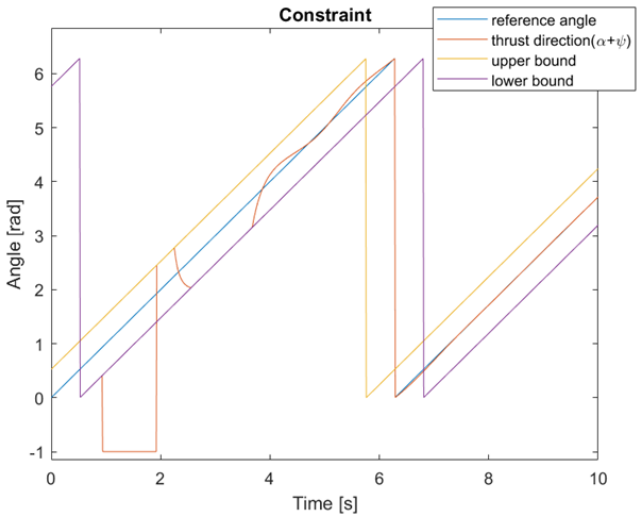


Fig. 23. The direction of the thrust (angle -1 represents 0 acceleration).

Judged from Figure 20—21, the trajectory generally follows the reference as a circle after some time. In Figure 23, we can see that the inputs do not diverge in two opposite directions. Actually, they become identical in the end. This is because that the number of the inputs is equal to the number of the outputs; State Drift can happen only in an actuated system. For a system with an equal number of inputs and outputs, State Drift will not happen.

There is an undesired behavior in this controller: the saturation in inputs.

The feedback linearization transfers the nonlinear system to a linear system (as what we have done in Equation (19)). So that the linear controller can be applied. This approach is valid and solid if and only if the constraint is not effective; hitting the bound (e.g., positive thrust constraint) invalidates the feedback linearization method.

The nonlinear property of the original system should be considered if the saturation happens. However, there lacks stability proof for a general linear time invariant system. Only several

categories of linear time variant systems with specific forms can be proved stable [19,20]. Seeking the stability proof for the rest linear time variant systems is still an open question.

Unfortunately, when saturation happens in this controller, Equation (20), with the circular reference in Figure 5, the linear time variant system lies in the category whose stability proof is not found. The detail of the deducing process is omitted in this paper, since the analysis of a nonlinear system is beyond this research.

The saturation can also be observed in Figure 23, the direction of the thrust. It can be seen that the input signals saturate at around 3 seconds.

6. CONCLUSIONS AND DISCUSSIONS

The results in the over-actuated state feedback linearization show that the system is stabilized. The State Drift happens in the controlled over-actuated system, although the stability is not affected in this system.

Applying the gait plan to this system, the number of the inputs becomes equal to the number of the outputs. Subsequently, the State Drift phenomenon disappears. The saturation happens in this control method, which affects the stability proof.

Further steps can be deducing the stability proof for the control methods in gait plan with saturation and different gait patterns development.

In the experiments in this research, State Drift seems to only affect the allocation of the inputs. While the stability seems not be affected by this phenomenon. In this section, we will see a higher order feedback linearization controller. This controller is to stabilize the system in Equation (2)—(3). The State Drift in this example shows that the stability is affected.

6.1. Fourth Derivative Feedback Linearization

Calculate the fourth derivative of the position. It can be computed by differentiating Equation (7). The result yields:

$$\begin{bmatrix} \dddot{x} \\ \dddot{y} \end{bmatrix} = \frac{1}{m} \cdot J_\psi \cdot \ddot{J}_\alpha \cdot J_\theta \cdot \begin{bmatrix} {\omega_1}^2 \\ {\omega_2}^2 \end{bmatrix} + \frac{2}{m} \cdot J_\psi \cdot \dot{J}_\alpha \cdot J_\theta \cdot \begin{bmatrix} 2\omega_1 \cdot \dot{\omega}_1 \\ 2\omega_2 \cdot \dot{\omega}_2 \end{bmatrix} + \frac{1}{m} \cdot J_\psi \cdot J_\alpha \cdot J_\theta \cdot \begin{bmatrix} 2{\dot{\omega}_1}^2 + 2\omega_1 \cdot \ddot{\omega}_1 \\ 2{\dot{\omega}_2}^2 + 2\omega_2 \cdot \ddot{\omega}_2 \end{bmatrix} + \frac{2}{m} \cdot \dot{J}_\psi \cdot \dot{J}_\alpha \cdot J_\theta \cdot \begin{bmatrix} {\omega_1}^2 \\ {\omega_2}^2 \end{bmatrix} + \frac{2}{m} \cdot \dot{J}_\psi \cdot J_\alpha \cdot J_\theta \cdot \begin{bmatrix} 2\omega_1 \cdot \dot{\omega}_1 \\ 2\omega_2 \cdot \dot{\omega}_2 \end{bmatrix} + \frac{1}{m} \cdot \ddot{J}_\psi \cdot J_\alpha \cdot J_\theta \cdot \begin{bmatrix} {\omega_1}^2 \\ {\omega_2}^2 \end{bmatrix} \tag{21}$$

where

$$J_\psi = \begin{bmatrix} \cos(\psi) & -\sin(\psi) \\ \sin(\psi) & \cos(\psi) \end{bmatrix} \tag{22}$$

$$J_\alpha = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \tag{23}$$

Equation (21) can also be written in Equation (24).

$$\begin{bmatrix} \dddot{x} \\ \dddot{y} \end{bmatrix} = M_P + \Delta_P \cdot \begin{bmatrix} \ddot{\omega}_1 \\ \ddot{\omega}_2 \\ \ddot{\alpha} \end{bmatrix} \tag{24}$$

where $\Delta_P \cdot \begin{bmatrix} \ddot{\omega}_1 \\ \ddot{\omega}_2 \\ \ddot{\alpha} \end{bmatrix}$ includes all the terms containing $\ddot{\omega}_1$, $\ddot{\omega}_2$, or $\ddot{\alpha}$. M_P includes all the remaining terms without containing $\ddot{\omega}_1$, $\ddot{\omega}_2$, or $\ddot{\alpha}$.

Specifically,

$$\Delta_P = [\Delta_{P_\omega}^{2\times 2} | \Delta_{P_\alpha}^{2\times 1}] \tag{25}$$

Where

$$\Delta_{P_\omega}^{2\times 2} = \frac{1}{m} \cdot J_\psi \cdot J_\alpha \cdot J_\theta \cdot \begin{bmatrix} 2\omega_1 & 0 \\ 0 & 2\omega_2 \end{bmatrix} \tag{26}$$

$$\Delta_{P_\alpha}^{2\times 1} = \frac{1}{m} \cdot J_\psi \cdot \begin{bmatrix} -\sin(\alpha) & -\cos(\alpha) \\ \cos(\alpha) & -\sin(\alpha) \end{bmatrix} \cdot J_\theta \cdot \begin{bmatrix} {\omega_1}^2 \\ {\omega_2}^2 \end{bmatrix} \tag{27}$$

Notice that when ω_1 and ω_2 are not both zero, we have

$$Rank(\Delta_{P_\omega}^{2\times 2}) = 2 \tag{28}$$

Equation (28) indicates that Δ_P has full row rank when ω_1 and ω_2 are not both zero. With this condition, $pinv(\Delta_P)$ can be calculated based on Equation (12).

Separating the inputs in Equation (24) yields Equation (29).

$$\begin{bmatrix} \ddot{\omega}_1 \\ \ddot{\omega}_2 \\ \ddot{\alpha} \end{bmatrix} = pinv(\Delta_P) \cdot \left(\begin{bmatrix} \dddot{x} \\ \dddot{y} \end{bmatrix} - M_P \right) \tag{29}$$

PD controllers are built based on Equation (29).

The reference and the initial condition are identical to the settings in the previous experiments.

They are initial $\omega_1 = 200$, initial $\omega_2 = 200$, and Figure 5.

The dynamic state error and the input squared are in Figure 24 and 25, respectively.

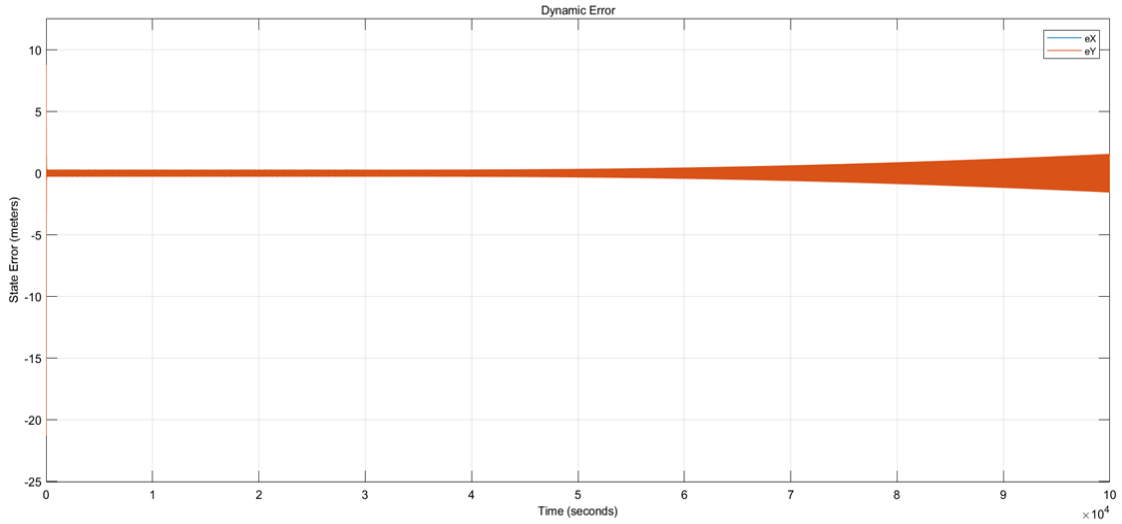


Fig. 24. Dynamic State Error.

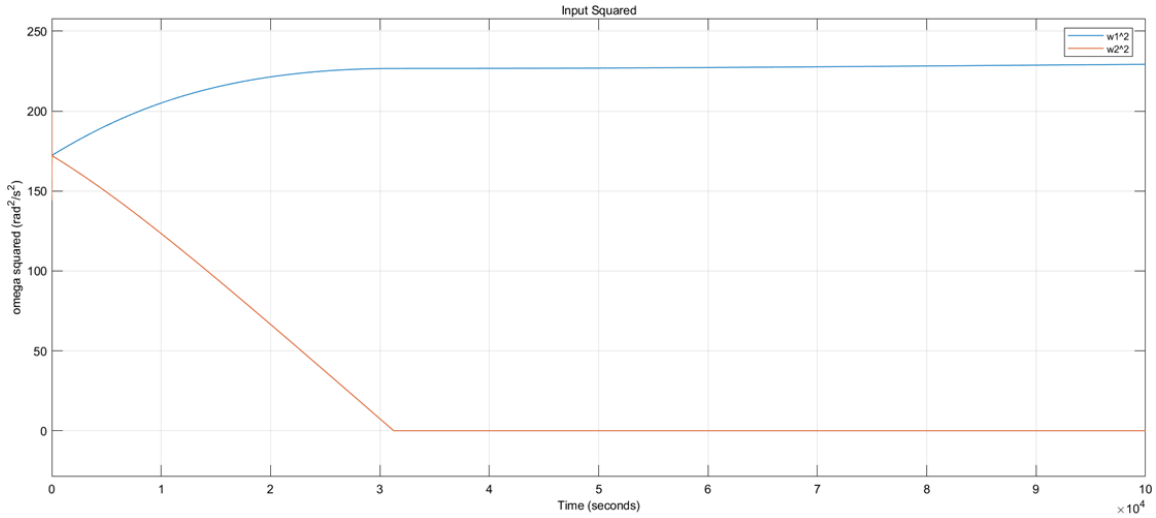


Fig. 25. Input Squared.

The dynamic state error (Figure 24) decreases to an acceptably small value (0.2 meter) at the beginning. While it increases after around 3×10^4 seconds. Finally, the dynamic state error diverges from 0, potentially making this system unstable.

State Drift causes this.

Judged from Figure 25, the State Drift happens so that the two inputs diverge in the two opposite directions. As we know from the previous experiments, it can not be a factor causing unstable.

However, after 3 seconds, the input ($\omega_2{}^2$) hit the positive constraint (negative thrust is not allowed), destroying the underlying stability proof. The consequence is what we see in Figure 24; unstable system with increasing dynamic state error is reported.

6.2. State Drift Phenomenon

So far, State Drift phenomenon has been thoroughly reported by controlling a fictional system (tilt vehicle). The reason can be multiple equilibrium states (inputs), which is explained in Figure 16—18.

However, the mathematical cause for pulling the system from one equilibrium to another is complicated and beyond the scope of this research.

For example, the following question is extremely hard to reply:

I understand that the system is stable even if State Drift happens, but why does State Drift happen?
The answer to this question can also be a challenged further step.

REFERENCES

[1] Taniguchi, T., & Sugeno, M. (2017). Trajectory tracking controls for non-holonomic systems using dynamic feedback linearization based on piecewise multi-linear models. *IAENG International Journal of Applied Mathematics*, *47*(3), 339-351.

[2] Mistler, V., Benallegue, A., & M'sirdi, N. K. (2001, September). Exact linearization and noninteracting control of a 4 rotors helicopter via dynamic feedback. In *Proceedings 10th IEEE international workshop on robot and human interactive communication. Roman 2001 (Cat. no. 01th8591)* (pp. 586-593). IEEE.

[3] Mokhtari, A., & Benallegue, A. (2004, April). Dynamic feedback controller of Euler angles and wind parameters estimation for a quadrotor unmanned aerial vehicle. In *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004* (Vol. 3, pp. 2359-2366). IEEE.

[4] Ryll, M., Bülthoff, H. H., & Giordano, P. R. (2012, May). Modeling and control of a quadrotor UAV with tilting propellers. In *2012 IEEE international conference on robotics and automation* (pp. 4606-4613). IEEE.

[5] Bouabdallah, S., Noth, A., & Siegwart, R. (2004, September). PID vs LQ control techniques applied to an indoor micro quadrotor. In *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)(IEEE Cat. No. 04CH37566)* (Vol. 3, pp. 2451-2456). IEEE.

[6] Lee, H., Kim, S., Ryan, T., & Kim, H. J. (2013, October). Backstepping control on se (3) of a micro quadrotor for stable trajectory tracking. In *2013 IEEE international conference on systems, man, and cybernetics* (pp. 4522-4527). IEEE.

[7] Lee, T., Leok, M., & McClamroch, N. H. (2010, December). Geometric tracking control of a quadrotor UAV on SE (3). In *49th IEEE conference on decision and control (CDC)* (pp. 5420-5425). IEEE.

[8] Xu, R., & Ozguner, U. (2006, December). Sliding mode control of a quadrotor helicopter. In *Proceedings of the 45th IEEE Conference on Decision and Control* (pp. 4957-4962). IEEE.

[9] Bouabdallah, S., & Siegwart, R. (2005, April). Backstepping and sliding-mode techniques applied to an indoor micro quadrotor. In *Proceedings of the 2005 IEEE international conference on robotics and automation* (pp. 2247-2252). IEEE.

[10] Das, A., Subbarao, K., & Lewis, F. (2008, September). Dynamic inversion of quadrotor with zero-dynamics stabilization. In *2008 IEEE International Conference on Control Applications* (pp. 1189-1194). IEEE.

[11] Voos, H. (2009, April). Nonlinear control of a quadrotor micro-UAV using feedback-linearization. In *2009 IEEE International Conference on Mechatronics* (pp. 1-6). IEEE.

[12] Shen, Z., & Tsuchiya, T. (2021). Singular Zone in Quadrotor Yaw-Position Feedback Linearization. *arXiv preprint arXiv:2110.07179*.

[13] Şenkul, F., & Altuğ, E. (2013, May). Modeling and control of a novel tilt—Roll rotor quadrotor UAV. In *2013 International Conference on Unmanned Aircraft Systems (ICUAS)* (pp. 1071-1076). IEEE.

[14] Nemati, A., Kumar, R., & Kumar, M. (2016, October). Stabilizing and control of tilting-rotor quadcopter in case of a propeller failure. In *Dynamic Systems and Control Conference* (Vol. 50695, p. V001T05A005). American Society of Mechanical Engineers.
[15] Rajappa, S., Ryll, M., Bülthoff, H. H., & Franchi, A. (2015, May). Modeling, control and design optimization for a fully-actuated hexarotor aerial vehicle with tilted propellers. In *2015 IEEE international conference on robotics and automation (ICRA)* (pp. 4006-4013). IEEE.
[16] Ryll, M., Bülthoff, H. H., & Giordano, P. R. (2014). A novel overactuated quadrotor unmanned aerial vehicle: Modeling, control, and experimental validation. *IEEE Transactions on Control Systems Technology*, *23*(2), 540-556.
[17] Elfeky, M., Elshafei, M., Saif, A. W. A., & Al-Malki, M. F. (2013, June). Quadrotor helicopter with tilting rotors: Modeling and simulation. In *2013 world congress on computer and information technology (WCCIT)* (pp. 1-5). IEEE.
[18] Zamani, A., Khorram, M., & Moosavian, S. A. A. (2011, October). Dynamics and stable gait planning of a quadruped robot. In *2011 11th International Conference on Control, Automation and Systems* (pp. 25-30). IEEE.
[19] Wu, F., & Grigoriadis, K. M. (1997, December). LPV systems with parameter-varying time delays. In *Proceedings of the 36th IEEE Conference on Decision and Control* (Vol. 2, pp. 966-971). IEEE.
[20] Zhou, B. (2016). On asymptotic stability of linear time-varying systems. *Automatica*, *68*, 266-276.

AUTHORS

Zhe Shen received his M.S. degree from Technical University of Denmark, Denmark, in 2019. In 2020, he joined The University of Tokyo, where he is currently a Ph.D. student of the Department of Aeronautics and Astronautics.

Takeshi Tsuchiya received his M.E. and Ph.D. degrees from The University of Tokyo, Japan, in 1997 and 2000, respectively. In 2002, he joined The University of Tokyo, where he is currently a Professor of Department of Aeronautics and Astronautics.

PIANO4PLAY: AN AUTOMATED PIANO TRANSCRIPTION AND KEYBOARD VISUALIZATION SYSTEM USING AI AND DEEP LEARNING TECHNIQUES

Jinge Liu[1] and Shuyu Wang[2]

[1]Portola High school, 1001 cadence, Irvine, CA 92618
[2]University of Minnesota, 2900 University Ave, Crookston, MN 56716

ABSTRACT

Piano keyboard visualization was very popular right now, but there are very few virtual piano keyboard visualizations right now [1]. I was using unity to show the virtual piano keyboard and then they can play piano pieces by themselves or play a recording online [2]. After that you can listen and see how the recording pieces play it on the visual keyboard to give them a clear idea about how the songs played on a keyboard2 [3]. For those who played by themselves it can let them heard and know also when the visual piano play for them, they can tell if they have offbeat playing or they missing not.

Piano4Play is an automated piano transcription and keyboard visualization system using AI and deep learning techniques. The user could upload a recorded piece of music, and our app would visualize the music on a digital piano keyboard. The user could see how the music is played visually in order to help piano beginners to see how the music will be played on piano in order to help them learn more quickly and easier, and advanced players could use the app to see whether they made any mistake when they are playing so they can get some improvement. Our app uses wav and MIDI files, repl, real-time database, google Collab and Unity.

KEYWORDS

3D Modeling, Machine Learning, Data Science.

1. INTRODUCTION

As we know that piano visualize was the visualize playing the piano, it must be the most vivid detail possible with the calm state of your mind and also can see the keyboard with the absolute clarity. The reason of making piano4play app was want to help the piano learner, people who already know how to play piano or the people who don't know how to play piano, etc. For the piano learners, you can just record a piece and then you will see how to play this and how hard is this piece. For people who already know how to play the piano is the way that when you practice by yourself, sometimes you don't know what you got wrong, what you need to improve more. When you record on this app you can see are you playing fast on someplace, did you miss a note that you didn't realize, did you forget the sharp or flat when you check your recording between your play and you're recording, etc. When the teacher is next to you they can tell you what you got wrong but when the teacher is not with you, you need to figure by yourself and need to improve when you practice, so come try this app. I want to make this app different from the other products intended to solve the same challenges. I also know that technology is advanced now,

David C. Wyld et al. (Eds): CMC, SOFT, MLT, SAI - 2022
pp. 19-26, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.120502

and there are many products similar to mine. However, my product is from my experience of playing the piano since I was a child, to decide how to help people who have just learned piano to understand some piano music, and how to help those students who are taking the exam to choose the correct music they want/suitable for. How to help those who don't know how to play piano music; what the difficulty is, help those who don't know where they are wrong when they practice piano, where they are listening fast and full, so that they can better understand how to play the piano. The later practice plays even better.

Due to the research of Piano keyboard Visualization I found one person who called Xiaoxiao made things about piano keyboard visualization called << Andante: Walking Figures on the Piano Keyboard to Visualize Musical Motion>> [5]. I knew that they chose to make Andante, which is the animated character walking on the piano keyboards to walking step by the physical key or rhythm [4]. They use step by human which is the most fundamental human rhymes to show the understanding of the human step by beat.This can help them more visualizations on the preliminary prototype when pianists practice, improvisation and when they make their own song. Another research I did due to the Piano keyboard visualization was called <<Real-time music visualization using responsive imagery>> by Robyn Taylor, this work was they turned the music into a visualization system. When performing live, according to the music played by the performer, different images corresponding to the played music are mapped out. Therefore, it can make people enjoy music more and let more people know that these two points can actually be combined. Due to those two research I knew there was no one using Unity to make the project. I want to use unity to show the different style and different types of the piano keyboard visualization, different experiences of anyone.

I'm using unity and python for the back serve. How it worked was that it was an app at first and the other app was the (). Then go to the unity front-end game module and then go to the python back-end server, after this there are two things one is ML algorithm another was all the storage.

In the beginning, was thinking about a few points, unity app which can record the audio and upload a .wav file to the Firebase storage; Google collab notebook which will convert the .wav file to the MIDI format and it can be uploaded back to the Firebase storage; The Repl.it program will be download the MIDI file and convert it into a list with all the notes data and upload all the data to the Firebase real-time database; when the navigated to the piano playback menu in the Unity app, the notes data will be downloaded and the song will be played back based on the notes from the database [6][7]. But after making the step and thinking more about it having something I want to add on there and make it more perfect to what I really want to be, the unity app can record the audio and send the request to the server which is my google Colab. And when it selects a song, it needs to fetch and use the server Url that is stored in the Firebase real-time database; also for right now wants to send requests to the http://ip_address/midi "sone" as a key and can be value .wav file which is called record audio. At least, want to check what missing and want to improve on it make it more be perfect, more interesting, the last few things were: google collab notebook will be converted to the .wav file to be the MIDI format also can be converted into a list with all the notes data and will be upload all data to the Firebase real-time database; when the navigated to the piano playback menu in mine Unity app, the note data will be downloaded and the song will be played back based on the notes from the database.

- Problems:

-Piano learners often face the challenges that they do not have a both visually and hearing view of the tunes that they are going to play and learn.
-The lack of a visual piana keyboard in the market.

- Procedures:

-Users could use Piano4Play to show the virtual piano keyboard and then they can play piano pieces by themselves or play a recording online.
-Users can see how the recorded pieces are played online visually.
-Functions: give users the idea of how to play the recorded pieces or help users to check their playing.
-Using Firebase real-time database to save the final file [8].

- Conclusion:

-According to the data that we collected from a survey, most beginners and advanced piano players would like to use this kind of app. The most common reason is they like visualizing piano playing on our own mobile devices.

The rest of the paper is organized as follows: Section 2 gives the details on the challenges that we met during the experiment and designing the sample; Section 3 focuses on the details of our solutions corresponding to the challenges that we mentioned in Section 2; Section 4 presents the relevant details about the experiment we did, following by presenting the related work in Section 5. Finally, Section 6 gives the conclusion remarks, as well as pointing out the future work of this project.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. Convert Piano Wav file to MIDI file and save it to Firebase Realtime database

First of all, I do not know how to convert the wav. file to the MIDI file. I was using the online tools that try to make those tools that I can not put into the server and also that I am not able to put in the code to run it. So that later on, I did the research and then found the python library, which I found very useful to help me improve my python programming skill [9]. Then I knew that this thing is capable of helping the project to convert the wav file to MIDI file. Although we have tried hard to use all resources we can find to solve the problem we have met, there is another problem right now. If we want to save the MIDI file as a database in our server, the size of this file requires a very heavy load of the server which is too big and expensive for us to afford and maintain. However, mongo and sql are not appropert, so last we decided to use Firebase real time database to save the final file.

2.2. How to make the file more fluently using the output and simulate the keyboards

I want to make the file more fluent by using the output and more simulating the keyboard. So first we tried to use 2d mode to demonstrate the data stream but none of them work well to what we think. So later we decided to use 3D unity software platform to simulate the virtual keyboard scenes the unity supported by c# programming language. Which I don't familiar with it, so I try to learned it by myself with the material that online, also found that C#, the programming language Community is very useful, I could find a lot of information there [10].

2.3. To choose a stable framework to build a web-service response data stream

To host the back-end, we need to choose a stable framework to build a web-service response data stream. Python flask and the Java spring are two options we have. I choose python flask since python is much easier to launch. For learning the python back-end, I have been learning for long time. Try to understand how I can do this to my app and how it can help for my app. Then we integrate all the code together into a main function in the python flask framework and run in a port to do the interaction with API calls. The API takes a Wav file as an input and returns a MIDI file as the output. Then after that, we launch the framework on the AWS server to stable response to all the API calls. But to use the AWS server we need to set up all the libraries and environments. The python version and dependence also matter.

3. SOLUTION

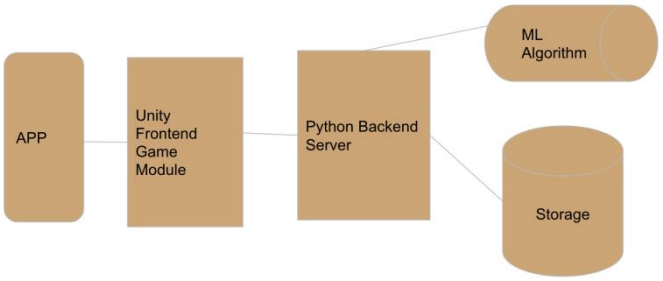


Figure 1. An overview of the project

Paino4Play is an automated piano transcription and keyboard visualization system using AI and deep learning techniques [11]. At the beginning, the unity app was used to record the audio, and then uploaded to the .wav file to Firebase for storage; google collab notebook will convert the .wav file into MIDI format, and upload it again to Firebase for storage; Repl [12]. It is a program to convert a MIDI file into a list with all the piano note data, and upload all the data to the Firebase real-time database; when these things are communicated to the piano in the unity app program it will play the menu, Compare the previously received note data with the notes in the existing database and then play the song. But at this time I wanted to add something here and make the whole thing more perfect and closer to the final look I had in mind. So it is intended that the unified application can record audio and send this request to the server which is my google collab. And when it wants to choose a song, it will need to get and can use those server URLs stored in the Firebase real-time database; also for right now wants to send requests to the http://ip_address/midi" sone" as a key and can be value .wav file which is called record audio. Finally, let google collab notebook be converted into .wav file into MIDI format or into a database list of all notes, etc., and all data can be uploaded to Firebase real-time database; when navigating to our Unity application The piano song play menu in the menu will automatically download the data about the notes and play the song according to the notes in the comparison database.

4. EXPERIMENT

4.1. Experiment 1

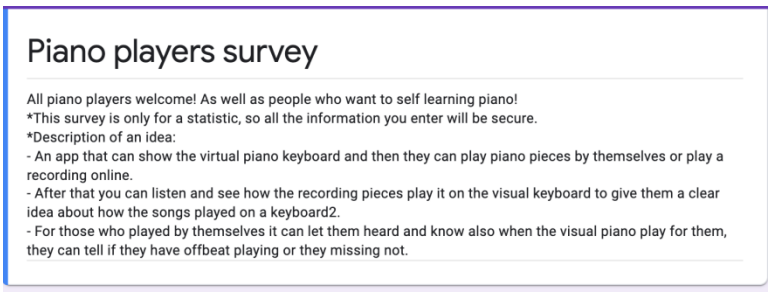


Figure 2. Survey result

First of all I made a survey called the piano players survey. In figure 2, it is what they will see when they open the survey. It says that there is an app that can show the virtual piano keyboards and then they can play piano pieces by themselves or play recording, after that they can listen and see how the recording pieces play it on the visual keyboard to give them a clear idea about how songs played on keyboard also. For those who played by themselves it can let them hear and know also when the virtual piano plays for them, they can tell if they have offbeat playing or missing out.

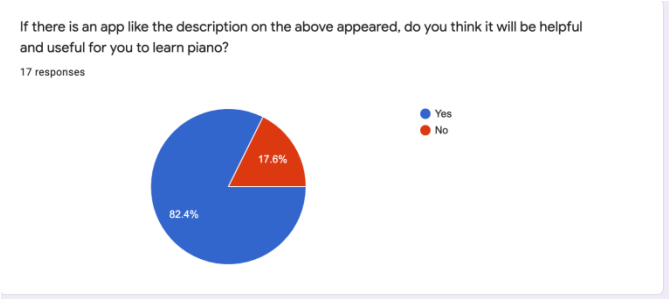
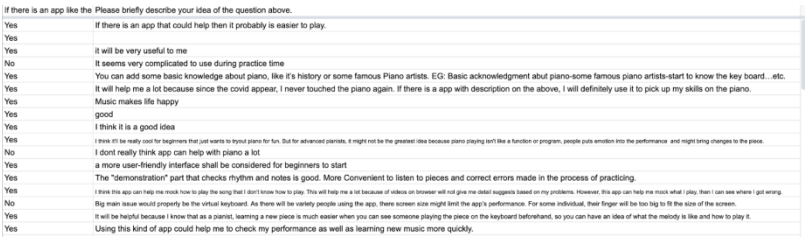


Figure 3. Survey chart

In figure 3 it was asking if this will be an app due to the description (figure 2), do they think it will be helpful or not. In that picture, it has the pie chart of the 17 responses. It shows that 82.4% of people think it was helpful, 17.6% think it is unhelpful.



If there is an app like the	Please briefly describe your idea of the question above.
Yes	If there is an app that could help then it probably is easier to play.
Yes	
Yes	it will be very useful to me
No	It seems very complicated to use during practice time
Yes	You can add some basic knowledge about piano, like it's history or some famous Piano artists. EG: Basic acknowledgment abut piano-some famous piano artists-start to know the key board…etc.
Yes	It will help me a lot because since the covid appear, I never touched the piano again. If there is a app with description on the above, I will definitely use it to pick up my skills on the piano.
Yes	Music makes life happy
Yes	good
Yes	I think it is a good idea
Yes	I think it'll be really cool for beginners that just wants to tryout piano for fun. But for advanced pianists, it might not be the greatest idea because piano playing isn't like a function or program, people puts emotion into the performance and might bring changes to the piece.
No	I dont really think app can help with piano a lot
Yes	a more user-friendly interface shall be considered for beginners to start
Yes	The "demonstration" part that checks rhythm and notes is good. More Convenient to listen to pieces and correct errors made in the process of practicing.
Yes	I think this app can help me mock how to play the song that I don't know how to play. This will help me a lot because of videos on browser will not give me detail suggests based on my problems. However, this app can help me mock what I play, then I can see where I got wrong.
No	Big main issue would properly be the virtual keyboard. As there will be variety people using the app, there screen size might limit the app's performance. For some individual, their finger will be too big to fit the size of the screen.
Yes	It will be helpful because I know that as a pianist, learning a new piece is much easier when you can see someone playing the piece on the keyboard beforehand, so you can have an idea of what the melody is like and how to play it.
Yes	Using this kind of app could help me to check my performance as well as learning new music more quickly.

Figure 4. Survey result 2

Figure 4 shows why people think it is helpful or not due to their answer on the previous question(figure 3).

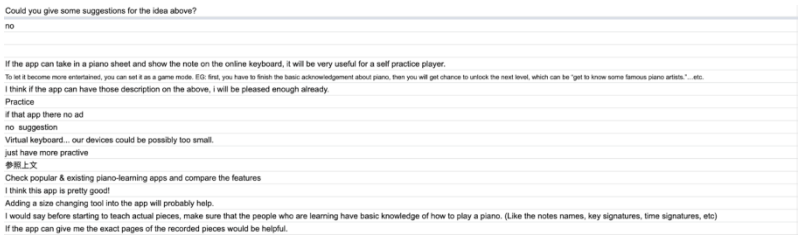


Figure 5. Survey result 3

Figure 5 shows that if this is an app, and if you think it is helpful or not, do you have any idea about it or do they want something specific due to the app description.

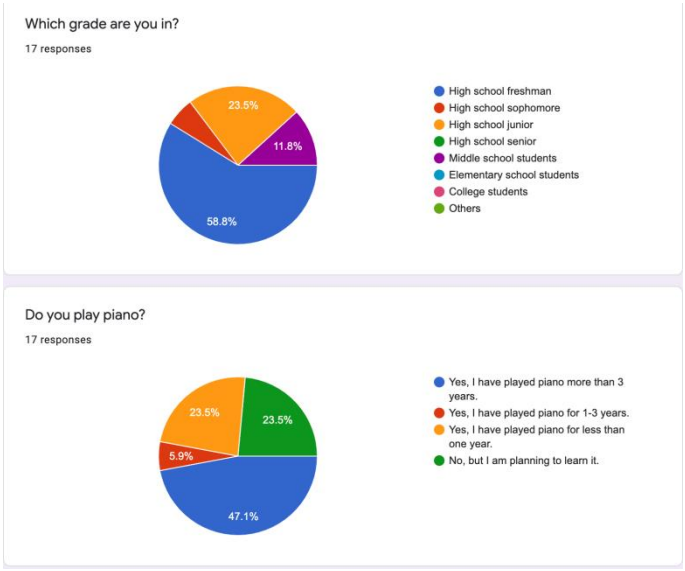


Figure 6. Survey result 4

For the figure 5, it shows which grade are the 17 respondents in, do they played piano before or not.

5. RELATED WORK

For their project, they use step by human which is the most fundamental human rhymes to show the understanding of the human step by beat [13].This can help them more visualizations on the preliminary prototype when pianists practice, improvisation and when they make their own song.

When performing live, according to the music played by the performer, different images, corresponding to the played music are mapped out.

When performing live, according to the music played by the performer, different images corresponding to the played music are mapped out [14]. Therefore, it can make people enjoy

music more and let more people know that these two points can actually be combined. Due to those two research I knew there was no one using Unity to make the project. I want to use unity to show the different style and different types of the piano keyboard visualization, different experiences of anyone.

Their spectrum analysis discovery uses an MDI file to do the piano visualization, which inspired me to use this skill in our projects.

In this article, they propose a new signal processing technique called spectrum analysis [15]. This can be visually demonstrated with a piano roll ten similar to chord music. Therefore Spectrum is defined here as the linear frequency transformation of the line segment of this number frequency. They experimented with multi-pitch signals containing sounds with different structural patterns, with varying degrees of parallel displacement superimposed on the more common spectral patterns in the logarithmic frequency domain or the overall shape of the multi-tone spectrum. They choose to convert to MIDI signals, which are crucial in the retrieval of music information.

Due to the research I knew that there is no such a project that uses Unity as a part. And unity can show the different style and different types of the piano keyboard visualization, different experiences of anyone.

6. CONCLUSIONS

This is an automated piano transcription and keyboard visualization system using AI and deep learning techniques. The user could upload a recorded piece of music, and the app would visualize the music on a digital piano keyboard. The user could see how the music is played visually in order to help piano beginners to see how the music will be played on piano in order to help them learn more quickly and easier, and advanced players could use the app to see whether they made any mistake when they are playing so they can get some improvement. Our app uses wav and MIDI files, repl, real-time database, google Collab and Unity. When doing this app, the main problem was: Can computers help piano players to visualize music, enable beginners to learn piano in a relatively easier way, and improve players' skills? For this problem I was trying to use the testing and the form to get the answer at the beginning, then I got the solution during the project time. Other problems about the project were: 1)How to Convert Piano Wav file to a MIDI file and save it to the Firebase Realtime database ? Solution for this was using the python library to convert wav and MIDI files. 2) How to save the MIDI file as a database in our server because the file is very big? Solution for this was using firebase real time database to save the final file. 3) How to make the file more fluently using the output and simulate the keyboards? Solution for this was 2D mode => cannot work well; 3D unity software: simulate the virtual keyboard scenes the unity supported by c# programming language. 4) The last question was How to host the backend using python or java? The solution was using python flask because it is so much easier to launch, API & AWS, also required for all the libraries and environments of the Python language. Due to all of that for future work I want to make the user interface and experience better like: slowing down playback for the beginner , scrubbing to certain parts of the playback, automatic recognition of errors, looping certain portions of playback. Also, due to the survey I mentioned before there are lots of people saying that adding the upload button for the music sheet when the 3D piano board is playing. Also, they want to add a metronome which can control the tempo or can make people more understand the beat.

REFERENCES

[1] Van Wijk, Jarke J. "The value of visualization." VIS 05. IEEE Visualization, 2005.. IEEE, 2005.
[2] Kang, Seongjae, Jaeyoon Kim, and Sung-eui Yoon. "Virtual piano using computer vision." arXiv preprint arXiv:1910.12539 (2019).
[3] Donison, Christopher. "Hand size vs the standard piano keyboard." Medical Problems of Performing Artists 15.3 (2000): 111-114.
[4] Bolton, Thaddeus L. "Rhythm." The american journal of psychology 6.2 (1894): 145-238.
[5] Shneiderman, Ben. "The eyes have it: A task by data type taxonomy for information visualizations." The craft of information visualization. Morgan Kaufmann, 2003. 364-371.
[6] Moore, F. Richard. "The dysfunctions of MIDI." Computer music journal 12.1 (1988): 19-28.
[7] Loy, Gareth. "Musicians make a standard: The MIDI phenomenon." Computer Music Journal 9.4 (1985): 8-26.
[8] Moroney, Laurence. "The firebase realtime database." The Definitive Guide to Firebase. Apress, Berkeley, CA, 2017. 51-71.
[9] Van Rossum, Guido, and Fred L. Drake. "Python library reference." (1995).
[10] Stefik, Andreas, and Stefan Hanenberg. "The programming language wars: Questions and responsibilities for the programming language community." Proceedings of the 2014 ACM International Symposium on New Ideas, New Paradigms, and Reflections on Programming & Software. 2014.
[11] Raphael, Christopher. "Automated Rhythm Transcription." ISMIR. Vol. 2001. 2001.
[12] Ellis, Kevin M., et al. "Write, execute, assess: Program synthesis with a repl." (2019).
[13] Xiao, Xiao, Basheer Tome, and Hiroshi Ishii. "Andante: Walking Figures on the Piano Keyboard to Visualize Musical Motion." NIME. 2014.
[14] Taylor, Robyn, Pierre Boulanger, and Daniel Torres. "Real-time music visualization using responsive imagery." 8th International Conference on Virtual Reality. 2006.
[15] Sagayama, Shigeki, et al. "Specmurt anasylis: A piano-roll-visualization of polyphonic music signal by deconvolution of log-frequency spectrum." ISCA Tutorial and Research Workshop (ITRW) on Statistical and Perceptual Audio Processing. 2004.

THE APPLICATION OF TECHNIQUES DERIVED FROM ARTIFICIAL INTELLIGENCE TO THE PREDICTION OF THE SOLVENCY OF BANK CUSTOMERS: CASE OF THE APPLICATION OF THE CART TYPE DECISION TREE (DT)

Karim Amzile and Rajaa Amzile

Faculty of Law, Economics and Social Sciences Agdal.
Mohammed V University of Rabat, Morocco

ABSTRACT

In this study we applied the CART-type Decision Tree (DT-CART) method derived from artificial intelligence technique to the prediction of the solvency of bank customers, for this we used historical data of bank customers. However we have adopted the process of Data Mining techniques, for this purpose we started with a data preprocessing in which we clean the data and we deleted all rows with outliers or missing values as well as rows with empty columns, then we fixed the variable to be explained (dependent or Target) and we also thought to eliminate all explanatory (independent) variables that are not significant using univariate analysis as well as the correlation matrix, then we applied our CART decision tree method using the SPSS tool.

After completing our process of building our model (AD-CART), we started the process of evaluating and testing the performance of our model, by which we found that the accuracy and precision of our model is 71%, so we calculated the error ratios, and we found that the error rate equal to 29%, this allowed us to conclude that our model at a fairly good level in terms of precision, predictability and very precisely in predicting the solvency of our banking customers.

KEYWORDS

Data Mining, Credit Risk, Bank, Decision Tree, Artificial Intelligence, Credit, risk.

1. INTRODUCTION

A Decision Tree (DT) is one of the methods of decision support, originating from Data Mining techniques. Its structure is based on a set of choices in the graphic form of a tree. The various possible decisions are at the end of the branches, and are reached according to decisions taken at each stage of the learning process. DT is one of the methods of supervised learning techniques, it allows to model a hierarchy of dichotomous tests and it also allows to predict the class of individuals in the case of a variable to be explained binary, such as the case of the solvency of bank customers.

A Decision Tree (DT) is one of the methods of decision support, originating from Data Mining techniques. Its structure is based on a set of choices in the graphic form of a tree. The various possible decisions are at the end of the branches, and are reached according to decisions taken at

David C. Wyld et al. (Eds): CMC, SOFT, MLT, SAI - 2022
pp. 27-36, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.120503

each stage of the learning process. DT is one of the methods of supervised learning techniques, it allows to model a hierarchy of dichotomous tests and it also allows to predict the class of individuals in the case of a variable to be explained binary, such as the case of the solvency of bank customers.

Each DT technique uses its own fractionation algorithms and fractionation measurements. Using well-known DT algorithms, the ID3, C5.0 and C&RT (CART) methods use impurity measurements to select the fractionation attribute and the fractionation value/s. While ID3 (Han et al 2000) uses information gain (IG[1]), C5.0 uses the information gain ratio (IGR[2]), and CART [3](Koh et al 2004) uses the Gini Index (GI[4]) for impurity measurements. For CHAID uses the KHI-2 or the F statistic to choose the division variable (Koh et al 2004).

2. LITERATURE REVIEW

In recent years, several sectors have adopted DM techniques to analyze dependent variables (Cheng and Leu 2011, Cheng et al 2015; Lee et al 2016). Although DM techniques have been applied to the analysis of banking data with the aim of predicting the behavior of decisive or explaining variables. Several methods and algorithms, such as neural networks, SVM and DT, have been used in DM to perform data classification and predictive modelling (Ryua et al 2007; Kim 2008).

DM techniques have six functions: association, grouping, classification, prediction, estimation and sequencing. Classification is a supervised analysis function and the most commonly used risk management function (Efraim et al, 2010). Several methods from this category of methods, among these analytical methods for classification and prediction.

DT has a structure that is easy to understand and explain. Thus, a tree diagram is often used to solve a series of classification and decision-making problems to classify many attributes and make predictions; for example, such diagrams have been applied in medical diagnosis and disease classification studies (Khan et al, 2009; Ture, Tokatli, & Kurt, 2009). A DT can also be used for classification studies in engineering and management (Syachrani & Chung, 2013; Alipour et al., 2017).

3. THE COMPONENTS OF A DECISION TREE

A DT is a hierarchical, triangular staircase construction consisting of several elements:

- Nodes: There are two categories of nodes:
  - ✓ Root node (NR): the first node in the tree, which concerns the first test.
  - ✓ Intermediate nodes (NI): nodes from the lower levels provide tests to better partition the individuals in our sample, for these nodes, they receive arrows in inputs and send arrows in outputs.

[1]Information Gain (GI) measures the amount of information that an explanatory variable gives us about the variable to explain. It tells us how important a descriptor is. Information Gain (IG) is used to decide the order of independent variables in the nodes of a decision tree.
[2]The information gain ratio is a relationship between the information gain and the intrinsic information.
[3]Classification and Régression Trees (Classification et arbre de régression)
[4]Is a statistical measure of the distribution of a variable in a population.

• Arrows: They indicate the direction of the path to be followed until the final decision. In our case we use dichotomous trees, that is, from each node, we will have two outgoing arrows (True/ False); one indicates the decision to be made or the next test if the condition in the node is true, and the other if not.

• Leaves: are nodes of the last levels of the tree, they contain the decision or class predicted for the individuals. For this kind of knot, there are only inward arrows for these leaves.

4. THE CONSTRUCTION OF AN AD OF TYPE CART:

A Decision Tree of type CART proposed by (Beriman et al, 1984) is a binary tree that helps in the decision of a dichotomous variable (Y) for an individual whose explanatory variables are known. Its construction is based on recursive partitioning of individuals using learning data. This partitioning is done by a succession of severed nodes using the Gini index.

To this end, to build a CART tree, we must answer these questions:

A. The determination of a criterion to select the best division among all those eligible for the different variables.
B. A rule to decide that a node is terminal.
C. The assignment of each sheet to one of the classes (for the variable to be explained qualitative) or to a value of the variable to be explained (for the quantitative case).

To answer the above questions the following steps are followed:

A. The division criterion is based on the definition of a heterogeneity function. The objective is to subdivide individuals into two segments which are the most homogeneous in the sense of the variable to be explained. The heterogeneity of a node is measured by a function that returns 2 types of results:

    a. Zero if, and only if, the node is homogeneous, that is, the individuals belong to the same modality of the variable to be explained.
    b. Maximum when Y values are equiprobable (Over-learning) or highly dispersed (Under-learning).

For our case, to choose the root variable or the segmentation variable, that is relative to the top of our decision tree, for this purpose we will use the index of Gini[5], this index adopted by the algorithm CART, measures the frequency for an individual to be misclassified by the node test. This will allow us to choose the variable that achieves the best separation of individuals, we will use a segmentation criterion that we calculate for the different learning variables.
To calculate the value of the Gini-index the following formula is used:

Équation 1

$$GI(D) = \sum_{i=1}^{k} p_i(1 - p_i)$$

With:

[5]The Gini Index is a variant value from 0 to 1, where 0 means perfect equality and 1 means perfect inequality

D: all data used
k: The number of modalities of the dependent variable (Solvable or Non-Solvable)
p_i: Represents the probability of modality i (creditworthy or non-financial) in the sheet after segmentation.

For our case we have 2 modalities so we can write:

Équation 2

$$GI(D) = p_1(1 - p_1) + p_2(1 - p_2)$$
$$GI(D) = p_1 - {p_1}^2 + p_2 - {p_2}^2$$

We have:

$$p_1 + p_2 = 1$$

Therefore:

Équation 3

$$GI(D) = 1 - {p_1}^2 - {p_2}^2 = 1 - \sum_{i=1}^{k} p_i^2$$

B. To decide in which node we will stop, although the growth of the tree must stop at a given node, which thus becomes terminal or leaf. While we will stop the segmentation of the node, when it is homogeneous or when there is no longer an admissible partition or if the number of observations it contains is less than a threshold value (between 1 and 5).
C. The assignment of sheets to one of the classes depends on the type of variable to be explained (Y),

• For quantitative case Y, each sheet has a value that is the average of the observations associated with that sheet.
• In the qualitative case, each sheet or terminal node is assigned to a class (Solvable / Non-Solvable) of Y by considering one of the following conditional modes:

- The class best represented in the node;
- The most likely posterior class in the Bayesian sense;

5. THE PREDICTION OF THE SOLVENCY OF A BANK'S CUSTOMERS USING THE CART-DT:

5.1. Pre-Data Processing

a. variable definitions

For our study, we used data from a bank's customers, our database consists of 3988 rows and 14 columns, among the variables used we have 13 explanatory variables and a single variable to explain which represents the solvency of customers.

To ensure a better modelling we thought about the codification of nominal qualitative variables. The table below illustrates the nature of each variable used and its values after coding:

Table 1: Coding of variables used

	NAME_CONTRACT_TYPE(Boolean)
Cash loans	1
Revolving loans	0
	CODE_GENDER (Boolean)
F	1
M	0
	FLAG_OWN_CAR (Boolean)
Y	1
N	0
	NAME_INCOME_TYPE (Qualitative)
State servant	1
Working	2
Commercial associate	3
Pensioner	4
	NAME_FAMILY_STATUS (Qualitative)
Married	1
Single / not married	2
Civil marriage	3
Separated	4
Widow	5
	NAME_HOUSING_TYPE (Qualitative)
House / apartment	1
With parents	2
Municipal apartment	3
Office apartment	4
Co-op apartment	5
Rented apartment	6
	NAME_EDUCATION_TYPE (Qualitative)
Higher education	1
Incomplete higher	2
Secondary / secondary special	3
Lower secondary	4

b. univariate analysis

Based on the result of the univariate analysis on SPSS, we obtained the following table 2:

Table 2 : univariate analysis

Vi	*L'intitulé de la variable*	*B*	*E.S.*	*Wald*	*ddl*	*Sig*
V_2	**NAME_CONTRACT_TYPE**	,429	,132	10,522	1	,001
V_3	**CODE_GENDER**	-1,604	,165	94,361	1	,000
V_4	**FLAG_OW0_CAR**	-,424	,085	24,811	1	,000
V_5	**CNT_CHILDREN**	,341	,126	7,298	1	,007
V_6	**AMT_INCOME_TOTAL**	,000	,000	39,407	1	,000
V_7	**AMT_CREDIT**	,000	,000	34,029	1	,000
V_8	**AMT_ANNUITY**	,000	,000	,435	1	,510
V_9	**AMT_GOODS_PRICE**	,000	,000	44,091	1	,000
V_{10}	**NAME_INCOME_TYPE**	-,375	,063	35,415	1	,000
V_{11}	**NAME_EDUCATION_TYPE**	,746	,038	375,557	1	,000
V_{12}	**NAME_FAMILY_STATUS**	-,017	,043	,151	1	,697
V_{13}	**NAME_HOUSING_TYPE**	,108	,044	6,167	1	,013
V_{14}	**CNT_FAM_MEMBERS**	-,263	,111	5,650	1	,017
	Constante	1,497	,354	17,887	1	,000

Based on the univariate analysis, it can be concluded that the variables V_8 (AMT_ANNUITY) and V_12 (NAME_FAMILY_STATUS) are not significant because of their meaning values that exceed 5%, however they must be eliminated in order to build a powerful model in terms of significance by keeping only the variables that have a significant probability.

However, we can also check for qualitative explanatory variables, if there are correlations between dependent variables, to this effect we draw the following correlation matrix (Table: 3):

Table 3 : correlations matrix

correlations											
	V2	***V3***	***V4***	***V5***	***V6***	***V7***	***V9***	***V10***	***V11***	***V13***	***V14***
V2	*1*	*-,023*	*-,034**	*-,004*	*,000*	*,227*	*,188*	*-,027*	*,053*	*-,003*	*,007*
V3	*-,023*	*1*	*-,093*	*,059*	*-,026*	*,063*	*,062*	*,001*	*-,049*	*-,037*	*,079*
V4	*-,034**	*-,093*	*1*	*,083*	*,179*	*,108*	*,116*	*,024*	*-,146*	*-,053*	*,109*
V5	*-,004*	*,05*	*,083*	*1*	*-,056*	*-,033*	*-,042*	*-,013*	*-,014*	*,014*	*,888*
V6	*,000*	*-,026*	*,179*	*-,056*	*1*	*,379*	*,388*	*,196*	*-,273*	*-,058*	*-,067*
V7	*,227*	*,063*	*,108*	*-,033*	*,379*	*1*	*,986***	*,106*	*-,197*	*-,072*	*,012*
V9	*,188***	*,062***	*,116***	*-,042***	*,388*	*,986***	*1*	*,108*	*-,215*	*-,073*	*,005*
V10	*-,027*	*,001*	*,024*	*-,013*	*,196*	*,106*	*,108*	*1*	*-,104*	*-,005*	*-,028*
V11	*,053*	*-,049*	*-,146*	*-,014*	*-,273*	*-,197*	*-,215*	*-,104*	*1*	*,074*	*-,010*
V13	*-,003*	*-,037*	*-,053*	*,014*	*-,058*	*-,072*	*-,073*	*-,005*	*,074*	*1*	*-,015*
V14	*,007*	*,079*	*,109*	*,888*	*-,067*	*,012*	*,005*	*-,028*	*-,010*	*-,015*	*1*

According to the above matrix, we can notice that there is a strong correlation between V_9 and V_7 so between V_14 and V_5, for this purpose we will proceed to eliminate a single variable in each strongly correlated couple, to ensure better predictability of the model. For this we will eliminate V_9 and V _14.

So, we will keep only the following variables:

Vi	*L'intitulé de la variable*
1	**NAME_CONTRACT_TYPE**
2	**CODE_GENDER**
3	**FLAG_OW0_CAR**
4	**CNT_CHILDREN**
5	**AMT_INCOME_TOTAL**
6	**AMT_CREDIT**
7	**NAME_INCOME_TYPE**
8	**NAME_EDUCATION_TYPE**
9	**NAME_HOUSING_TYPE**

6. THE GRAPHICAL REPRESENTATION OF THE AD CART

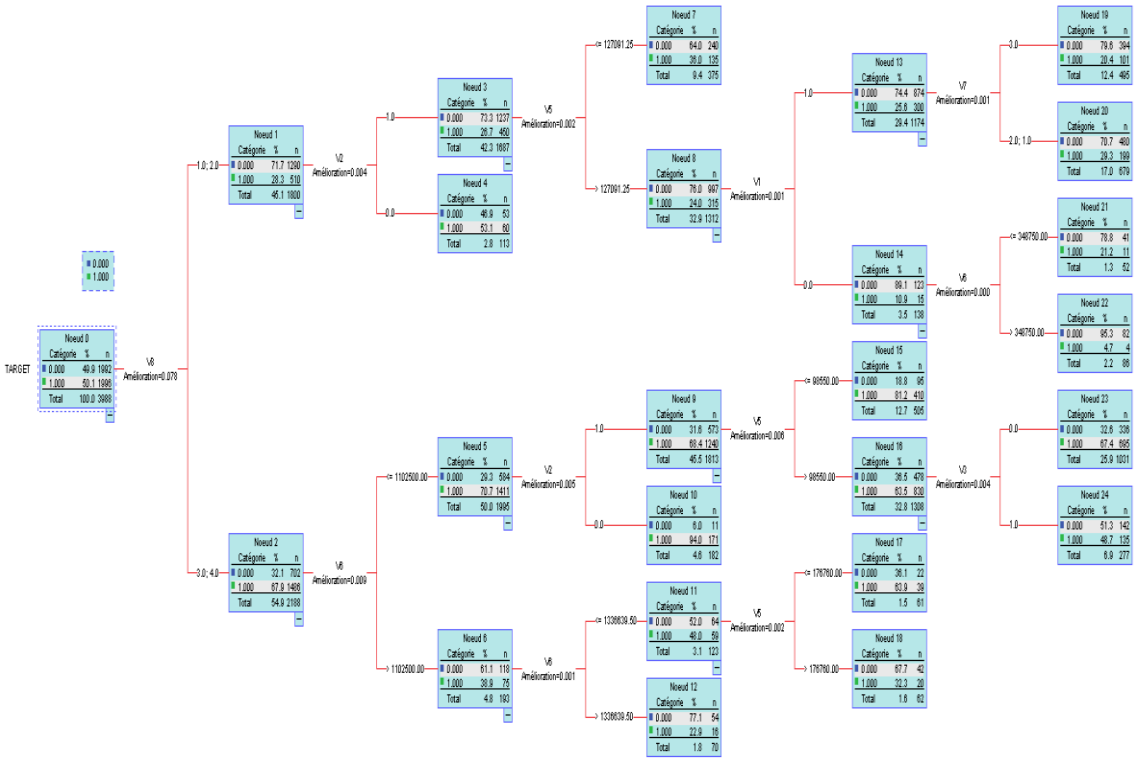


Figure 1 : DT- CART (SPSS / Source : Author)

7. ASSESSMENT OF OUR MODEL

7.1. Confusion Matrix

Classification			
Observations	Prediction		
	0	1	Pourcentage correct
0	1475	517	74,0%
1	621	1375	68,9%
Pourcentage global	52,6%	47,4%	71,5%
Method : CART Variable dependent : TARGET			

True Positives (VP): 1 correctly predicted VP = 1375
True Negatives (VN): 0 correctly predicted VN = 1475
False Negatives (FN): 0 incorrectly predicted FN = 517
False Positives (FP): 1 incorrectly predicted FP = 621
Sum = 3988

7.2. Measurement of error rates

We will use three error ratios:

$$e_1 = \frac{number\ of\ observations(1)\ classified(0)}{number\ of\ obs(1)}$$

$$e_2 = \frac{number\ of\ obs(0)\ classified\ (1)}{number\ of\ obs(1)}$$

$$e_3 = \frac{number\ of\ obs(1) classified(0) + number\ of\ obs(0)\ classified\ (1)}{number\ of\ obs}$$

To calculate these 3 types of errors we must build our confusion matrix:

So, we can now calculate the errors:

$$e_1^{cart} = \frac{621}{1996}\ e_2^{cart} = \frac{517}{1992}\ e_3^{cart} = \frac{1138}{3988}$$

	e_1	e_2	e_3
AD-Cart	0,31112224	0,25953815	0,285356

So according to the above matrix we can calculate the following metrics:

$$Accuracy = \frac{VP + VN}{Somme} = \frac{1375 + 1475}{3988} = 0.71$$

$$Sensibilité = \frac{VP}{VP + FP} = \frac{1375}{1375 + 621} = 0.69$$

$$Spécificité = \frac{VN}{VP + FP} = \frac{1475}{1375 + 621} = 0.74$$

7.3. the ROC curves

Area Under the curve				
Zone	Erreur Std	Signification	Confidence interval 95% asymptotique	
			Lower Bound	Upper bound
AUC = 0,715	0,008	0,000 << 0.05	0,698	0,731

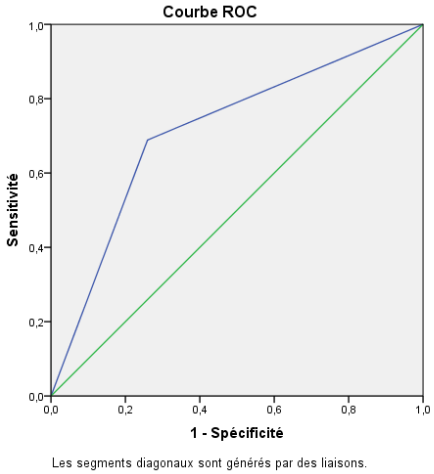


Figure 2 : ROC Curve

8. OUTCOME AND DISCUSSION

According to the previous illustrations of the evaluation results of our CART-type DT model, we were able to obtain a model with a precision of 0.71 which is quite good and we also noticed this precision in the graphical representation of the ROC curve, of which we obtained an area under the curve equal to $AUC = 0.715$. However, our model according to the values of $specificity = 0.74$ as well as $Sensitivity = 0.69$ reflects that our model is better performing in the classification of non-bankable clients than creditworthy clients.

9. CONCLUSION

After applying DT-CART in our study, we obtained satisfactory results; in terms of precision, as a result our model achieved an accuracy of 71%, therefore the error rate e_3 does not exceed 29%, and the ROC curve reflects a fairly good percentage of precision. Therefore, we can say that we have been able to build a model of new machine learning methods, with a fairly good performance and predictability level.

By way of conclusion, we can say that the techniques of Data Mining derived from artificial intelligence, can be used to solve problems of management of banking risks, to this end, we recommend exploiting this type of technique in the various types of financial risks, in order to reveal even more the opportunity offered by these methods to financial institutions and specifically to banks.

REFERENCES

[1] C.-L. Lin et C.-L. Fan, « Evaluation of CART, CHAID, and QUEST algorithms: a case study of construction defects in Taiwan », *Journal of Asian Architecture and Building Engineering*, vol. 18, nº 6, p. 539-553, nov. 2019, doi: 10.1080/13467581.2019.1696203.

[2] M.-A. Poursat, « Apprentissage statistique : Arbres de décision », p. 17.

[3] Y. Sahin et E. Duman, « Detecting Credit Card Fraud by Decision Trees and Support Vector Machines », *Hong Kong*, p. 6, 2011.

[4] Rachad EL MROUJI, « Elaboration d'un modèle de scoring appliqué à la microfinance.pdf », INSEA, 2020.

[5] C. Chesneau, « Introduction aux arbres de décision (de type CART) », p. 55.

[6] C. Decaestecker et M. Saerens, « Les arbres de décision (decision trees) », p. 22.

[7] Ricco RAKOTOMALALA, « rakotomalala-33-tutorial.pdf », *Revue MODULAD*, vol. 33, p. 25, 2005.

[8] S. Pang et J. Yuan, « WT Model & Applications in Loan Platform Customer Default Prediction Based on Decision Tree Algorithms », in *Intelligent Computing Theories and Application*, vol. 10954, D.-S. Huang, V. Bevilacqua, P. Premaratne, et P. Gupta, Éd. Cham: Springer International Publishing, 2018, p. 359-371. doi: 10.1007/978-3-319-95930-6_33.

AUTHORS

Karim AMZILE (PhD, Student)
Faculty of Law, Economics and Social Sciences Agdal.
Mohammed V University of Rabat, Morocco

Rajaa AMZILE (PhD, Professor)
Faculty of Law, Economics and Social Sciences Agdal.
Mohammed V University of Rabat, Morocco

WEBREVIEW: AN INTELLIGENT CLASSIFICATION PLATFORM TO AUTOMATE THE EVALUATION AND RANKING OF WEBSITE QUALITY AND USABILITY USING ARTIFICIAL INTELLIGENCE AND WEB SCRAPING TECHNIQUES

Darren Xu[1], Dexter Xu[1] and Ang Li[2]

[1]Redlands High School, 840 E Citrus Ave, Redlands, CA 92374, USA
[2]California State University, Long Beach, 1250 Bellflower Blvd, Long Beach, CA 90840, USA

ABSTRACT

Paywalls are a staple of the internet and seen in a vast amount of websites [1]. Encountering a paywall is always annoying, whether you're doing work for school or just trying to catch up on the latest news [2]. To eliminate this annoyance we have created Wall Breaker, a google extension with the primary task of bypassing any paywall using a variety of methods [3]. Our extension uses methods such as opening the website in an incognito tab or acting as a new user when clicking on a link. Although not the first of its kind, our extension is truly unique in the methods and techniques used. The popup used is easy to use and simple to look at, providing the best user experience. Wall Breaker will work on most websites, both popular and lesser known ones. It makes no distinction between certain types of websites and the methods can be used on any page. While Wall Breaker might not work on every website those are few and far between.

KEYWORDS

Web Scraping techniques, Google, paywall.

1. INTRODUCTION

Web-browsing and surfing the web have become a highly prevalent part of all aspects of society [4]. Many people wander into the realm of the internet to browse a variety of websites daily. The use of websites is integral and goes hand in hand with the internet. Whether their purpose is for research, entertainment, or other interests they have undoubtedly crossed paths with a paywall. Many websites hinder people's access to their website by requiring some form of payment, login, and or agreement. This is known as a paywall. Popular websites might use paywalls to restrict the users entry and cater towards those that have some kind of subscription. These confining stipulations deter possible enthusiasts and might cause annoyance in those who only wish to view the website a few times and do not want to sign up for a subscription of any kind. Getting around paywalls such as this opens up more information for internet users and improves the overall internet experience. It benefits the consumer by saving time, headache, and increasing the overall spread of knowledge and information to everyone. The ability to bypass a paywall is invaluable in order to increase access to information that otherwise would be barred from the normal viewer.

David C. Wyld et al. (Eds): CMC, SOFT, MLT, SAI - 2022
pp. 37-44, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.120504

The benefits of our extension, WallBreaker, opens up many different websites and minimizes the annoyance of everyday browsing. Because of how ingrained web-browsing is in our daily lives, any way to improve upon the now common activity is important.

Any internet browser has most likely encountered a paywall [5]. While certainly annoying, there are ways to get past one. A common use of paywalls is letting the user access a specific amount of pages and once that number is reached, activate the paywall preventing the user from reaching any new pages. While there are some manual ways to get by this such as using a friend's account or making a new one this is far from optimal. A much more efficient way to bypass a paywall is to do it automatically by using internal tools such as browser extensions. Our extension is not the first to attempt to fix this problem. There have been several different techniques and methods used by others in order to bypass paywalls. A popular method to bypass a paywall is to search the internet for a free version of that particular website by examining the contents of the page and searching the internet for a copy. However a common problem with this method is that oftentimes, a free copy is unable to be found on the internet. These limitations limit the amount of paywalls able to be bypassed. For example, a lesser known website with a paywall will most likely not have a free copy and even popular news articles and websites might not have any free versions available online. Most of the time free versions of websites with paywalls will be limited to scholarly articles. These methods also rely on already established databases that hold free versions of these websites.

We used a number of methods to achieve our goal of bypassing paywalls. The first step to achieving this goal is by making it easily available, we decided to create this project in the form of a google chrome extension. This extension has a brilliant UI designed to customize the user's needs. To add on, we created multiple options to make sure we can bypass all paywalls on all websites. Some examples include opening a private browser when hitting a paywall or disguising the user as a google bot beforehand. The functionality of each websites' paywall differs which results in the need for differing methods to bypass these obstacles. Some for example do not allow google bots but may allow social media links. We therefore have a method that disguises the user as coming from a social media platform to trick the website into allowing access to their content without signing up or paying for an account. Some websites' paywalls might be harder than others to traverse through. The inner workings of websites all over the web varies and their paywalls all need different requirements to bypass. This is why we required so many different solutions. Our solution has the strength of accessing the desired data directly. Other solutions such as unpaywall, while still a strong solution, reroute you to a different location where the data may possibly be stored. Our extension is also free to use, unlike some of the other solutions currently out there.. Due to these numerous features, we believe that our creation is invaluable while surfing the web. Our project is important to minimize the frustration of online browsing [6].

To prove our results we went to a series of websites and tested if we would be met with a paywall. Certain websites such as medium or wall street journal, contained paywalls with which we could run tests. To reliably test the results of our extension we tested the code manually, checking a selection of sites that offer a different variety in the types of paywalls that are used. For example, we tried our code against medium.com to test the functionality of the code. At each turn, our code reliably blocked the paywall. Our code has numerous working features, such as personally selecting certain websites to block or unblock. We tested this by using it on a variety of websites. Another feature is the ability to adapt to all kinds of paywalls by using a diverse set of methods built for a multitude of paywalls. These features are inclusive to all websites. Basically our extension is built for user customization and each function is able to work independently and as a whole. Ultimately our code works together to provide a better experience surfing the internet.

The rest of the paper is organized as follows: Section 2 gives the details on the challenges that we met during back-end and front-end coding process; Section 3 focuses on the details of our solutions corresponding to the challenges that we mentioned in Section 2; Section 4 presents the relevant details about the project we did, following by presenting the related work in Section 5. Finally, Section 6 gives the conclusion remarks, as well as pointing out the future work of this project.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. Creating an easy to use UI inside the pop-up

Creating an easy to use UI inside the pop-up was one of the main challenges when building our extension [7]. We wanted to create a pop-up that the average user could understand and use without difficulty [8]. As Wall Breaker has multiple methods to bypass paywalls we also needed a way to easily explain the different methods to bypass different types of paywalls. We decided to incorporate elements such as sliders and a simple design to create our pop-up.

2.2. Finding methods to bypass paywall

As each paywall is different it makes sense to have multiple methods to bypass them. Some of these methods only work for specific circumstances so putting on the solutions in one location was necessary but it proved to be a challenge. We had to create a script that had all the methods in one palace in order to integrate it all in the chrome extension. We used sliders to activate the methods. Making sure to save user settings and database info was also a challenge due to the fact that the sliders naturally reset each time the popup is opened. We thus had to create a series of code to make sure the settings changed.

2.3. Testing code

One of the more difficult parts of the code was actually testing it. Due to the nature of the extension it is basically impossible to see if the code works until we actually finish coding it, meaning that if there is an error it will most likely be unknown until we actually attempt to run the code. Chrome API communication between scripts leads to limitations in certain background tasks, making the process of building the extension much more restricted and difficult.

3. SOLUTION

There are many steps and components that make up the functionality of our project [9]. When the user first opens the extension they are greeted with an easy to use, and easy to understand user interface. This interface requests for four different types of inputs, which allow the user to customize their predilections and personalize their time online [10]. With one click you can personalize your online experience by excluding or including certain websites. Once the input is entered it is saved in a database that stores the user's preferences. The preferences are frequently updated with any new changes to the inputs that are made. To bypass the paywall our program uses a content script and a background script which work together in order to avoid the paywalls. The purpose of the content script is to find the best way to traverse through a paywall by going through numerous different strategies that might be implemented in passing the paywall. To elaborate, the content script finds what kind of website is opened-because every website utilizes a paywall differently- and observes the best solution in order to dodge the paywall. The content

script is constantly being updated with each click into a new website, Meanwhile the background script is refreshed only when the extension first runs. In order to continually refresh the background script, the background script is directly linked to the change of inputs from the user. The extension then returns the most optimal solution to bypass the paywall while simultaneously saving the user's settings into a database. Based on the two scripts and the input the extension customizes the user's experience on the web. Each of these individual mechanisms helps to run and produce the end result.

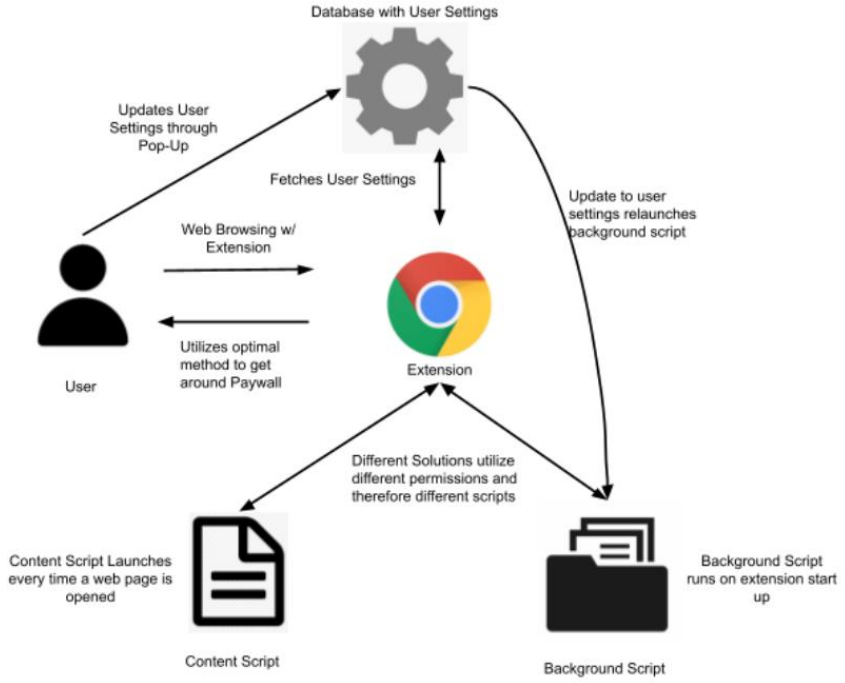


Figure 1. An overview of the project

Google Chrome Extensions are programs that change the browser in some way, whether it's the way Google looks or the way that the search bar works. Our extension manipulates the source code of websites to bypass paywalls. Wall Breaker uses multiple scripts that communicate and rely on each other to work. The content script checks the information of a website and then gives that information to other scripts to execute the code. The background script saves the user preferences on whether the sliders remain enabled or disabled. The manifest saves information about the extension like the name and icon used. Our extension uses four methods to bypass paywalls. These methods can be switched on and off with sliders. Multiple methods are used as some might not work on specific websites. Many websites let you view their page a couple of times before enabling a paywall and making you create an account. One of our methods automatically opens the website in an incognito page, making it seem as if one is visiting the website for the first time. Some websites disable the private browsing tab however so we also have a method that disables cookies on the website and one that modifies the HTTP headers to make it seem as if one is joining from a social media site. To add on, we also have a method that takes advantage of a Google bot where it changes the website request to make the user look like a GoogleBot. All these methods are unique in their own ways and help the user in bypassing paywalls on different websites in the form of a Google Extension.

4. EXPERIMENT

4.1. Experiment 1

In order to verify that our solution can effectively solve problems at different levels and have good user feedback, we decided to select multiple experimental groups and comparison groups for several experiments. For the first experiment, we want to prove that our solution works stable and continuously, so we choose a group size of 20 different IP Address in 5 different country The goal of the first experiment is to verify if the paywall detection works good for different IP address Through sampling 5 groups of users of different IP address and letting them try the same website 10 times. Results are collected by statistics of the total time that the paywall website is detected. Experiments have shown that all IP in different IP groups show a high rate of detection. IP form North American has the most high rates, IP form China has the most Low rates This experiment could explain that the IP address do have a obvious impact on the detection results The experiment graph shows below:

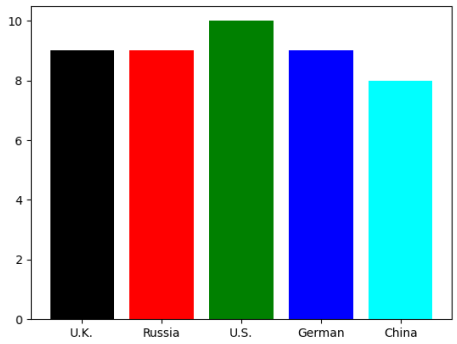


Figure 2. Survey result

4.2. Experiment 2

A good user experience is as important as a good product. So a perfect solution should have excellent user experience feedback. In order to prove that our solution has the best user feedback, we specially designed a user experience questionnaire We statistics the feedback result from 100 users, we divide those users into three different groups. The first group of users spend more time on the video games, the second group of users spend more time on the research and reading, the third group of users spend more time on working The goal of the first experiment is to verify high feedback scores shows high performance We collect the feedback scores form these 3 different group of users and analyze it. Experiments have shown that users who play games more give the highest result feedback to our app. Which may because of paywall link appears more in the game searching The experiment graph shows below:

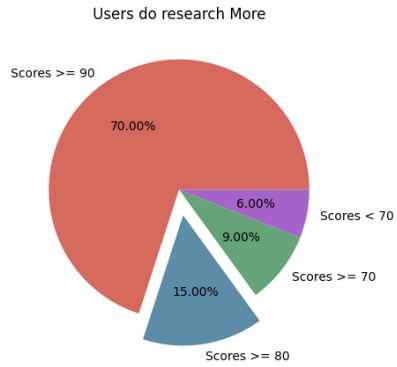


Figure 3. Survey result 2

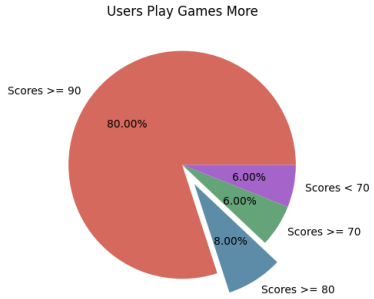


Figure 4. Survey result 3

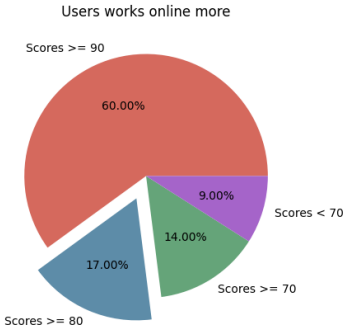


Figure 5. Survey result 4

5. RELATED WORK

Schultz and Azadbakht have a browser extension that allows the user to use “tools” that search for an open, free version of a website [11]. When encountering a paywall there are four options that show up on the screen that link to other websites where the user can read the article for free. Similar to our extension, the project extension uses tools that users click and use to gain access to websites. However, their extension relies on other already established websites that have databases of free versions of websites. Our extension actually manipulates the code of a website

and has a variety of different methods to bypass the paywall. For example, a lesser known website with a paywall might not have a free version but our extension can instead use a variety of methods to actually manipulate the code and gain access to the website.

Unpaywall is a widget that when encountering a paywall searches the internet for a free version of the website [13]. When encountering a paywall a tab will show up on the side of a screen that links to a free version or is grayed out, meaning it could not find a free version. Unpaywall is similar to our extension in that the user can interact with it. Of course, like the previous related work Unpaywall has restrictions and limitations of what it can do when encountering a paywall. Unpaywall searches for free versions of the website on the internet but if no such website is found the user has no choice but to find a different website or pay for it.

Libkey Nomad is a browser extension that finds free and full-text versions of journal articles [12]. Nomad examines the contents of an article page and searches their databases for a PDF of the article. Depending on if a PDF is found or not Nomad has several options for the user. For example, if a PDF is found the user can download the PDF and if a PDF is not found the user will be linked to a website where they can enable interlibrary loan. Libkey Nomad is limited to mostly scholarly articles but can help in accessing library subscriptions. Libkey Nomad is great for finding articles that need a library subscription to access. Our extension can access most websites and bypass the paywalls but might have trouble getting pass a subscription.

6. CONCLUSIONS

Encountering a paywall is always irritating to deal with especially when trying to do something important, like school work. Although bypassing a paywall manually is possible in some cases, such as creating a new account for a website, it is far from optimal and efficient. To effectively avoid encountering paywalls we have created Wall Breaker, a google extension that bypasses paywalls on the internet [14]. We used a variety of ways and methods to bypass paywalls. We added these methods into an easy to use popup that users can interact with. Users can turn these methods on and off and the popup will save their choices. One of the methods we use opens a page in the incognito tab. If the user does not want this to happen they can turn off this method permanently by clicking the slider. We used a variety of methods since the effectiveness of a method depends on the website. A method could work perfectly fine on a website but be completely ineffective on another [15]. The creation of multiple methods and functions lower the chance of running into a website with a paywall that can not be bypassed. The popup we created is simplistic to make sure the user can easily understand and use it. The popup also hides itself until the user interacts with it by clicking the icon on the taskbar. This makes sure that our extension is not distracting to the user. Wall Breaker is extremely effective in doing its task and makes sure that the user receives the best possible experience.

While our extension is extremely flexible there are some limitations associated with it. Our extension works on most websites but there might be a few occasions where none of our methods will work and there is no way to get by the paywall. Although our popup is easy to use it could be annoying to turn the sliders on and off depending on the website and trying to figure out which option is the best. We have multiple methods but the effectiveness of these methods vary from website to website and some can be annoying to leave on permanently such as the one that opens the website into an incognito page.

Most of the problems with our extension can be fixed by changing or adding new functions and code. We can change the code so that when encountering a paywall, a method will be automatically used without the need for user input. We can also add new and more methods to make it even less likely for there to be a paywall that can not be bypassed.

REFERENCES

[1] Estok, David. "Paywalls." Journal of Professional Communication (2012).
[2] Chiou, Lesley, and Catherine Tucker. "Paywalls and the demand for news." Information Economics and Policy 25.2 (2013): 61-69.
[3] Carlini, Nicholas, Adrienne Porter Felt, and David Wagner. "An evaluation of the google chrome extension security architecture." 21st {USENIX} Security Symposium ({USENIX} Security 12). 2012.
[4] Lieberman, Henry, Neil W. Van Dyke, and Adrian S. Vivacqua. "Let's browse: a collaborative Web browsing agent." Proceedings of the 4th international conference on Intelligent user interfaces. 1998.
[5] Muglerab, Emily, et al. "Control of an internet browser using the P300 event-related potential." International Journal of Bioelectromagnetism 10.1 (2008): 56-63.
[6] Kapoor, Ashish, Winslow Burleson, and Rosalind W. Picard. "Automatic prediction of frustration." International journal of human-computer studies 65.8 (2007): 724-736.
[7] Hoiem, Derek, Alexei A. Efros, and Martial Hebert. "Automatic photo pop-up." ACM SIGGRAPH 2005 Papers. 2005. 577-584.
[8] Greco, JoAnn. "From pop-up to permanent." Planning 78.9 (2012): 14-18.
[9] Wold, Svante, Kim Esbensen, and Paul Geladi. "Principal component analysis." Chemometrics and intelligent laboratory systems 2.1-3 (1987): 37-52.
[10] Reinhart, Tanya. "Interface strategies." OTS working papers in linguistics (1995).
[11] Azadbakht, Elena; Schultz, Teresa.Information Technology and Libraries (Online); Chicago Vol. 39, Iss. 2, (Jun 2020): 1-13. DOI:10.6017/ital.v39i2.12041
[12] Hoy, Matthew B. “LibKey Nomad.” Journal of the Medical Library Association : JMLA vol. 108,4 (2020): 672-674. doi:10.5195/jmla.2020.1017
[13] Chawla, Dalmeet Singh. "Unpaywall finds free versions of paywalled papers." Nature News (2017).
[14] Pattabhiramaiah, Adithya, S. Sriram, and Puneet Manchanda. "Paywalls: Monetizing online content." Journal of marketing 83.2 (2019): 19-36.
[15] Mintzer, Fred, Gordon W. Braudaway, and Minerva M. Yeung. "Effective and ineffective digital watermarks." Proceedings of International Conference on Image Processing. Vol. 3. IEEE, 1997.

BTF PREDICTION MODEL USING UNSUPERVISED LEARNING

Soichiro Kimura[1], Kensuke Tobitani[2] and Noriko Nagata[1]

[1]Kwansei Gakuin University, Hyogo, Japan
[2]University of Nagasaki, Nagasaki, Japan

ABSTRACT

The impressions evoked by textures are called affective textures, and are considered to be important in evaluating and judging the quality of an object. And, technologies for understanding and controlling sensory textures are needed in product design. In this study, we propose a BTF prediction method using DNN as a first attempt to generate textures based on affective texture recognition. The method uses a series of continuously varying viewpoint angles of a texture image as the input signal. This method enables the generation of texture images with continuously changing angles. We tested the validity of the proposed method by using textile, wood and paper. The results show that the proposed method is effective for predicting diffuse reflection optical properties and irregular and regular patterns.

KEYWORDS

PredNet, Machine Learning, BTF, Affective texture.

1. INTRODUCTION

The impression evoked by the surface properties (texture) of a material is called "affective texture," and is considered to be important in evaluating and judging whether an object is good or bad [1,2]. Therefore, technologies to understand and control affective textures are needed in product design. The reflectance characteristics of an object's surface, which are closely related to the affective texture, are represented by the Bidirectional Reflectance Distribution Function (BRDF) and the Bidirectional Texture Function (BTF) [3,4]. Because of its accuracy, it is widely used in texture representation. Conventional BRDF and BTF acquisition methods, however, require measurement with dedicated equipment, which is time and facility consuming and resource intensive. For this reason, Researchers have conducted studies on prediction and interpolation of BRDFs and BTFs [5].

On the other hand, in the research related to deep learning, which has been rapidly developing in recent years, methods related to image generation, such as Generative Adversarial Networks (GANs), have been used in various applications, such as image resolution enhancement and image style conversion [6, 7]. These methods have also attracted attention in texture representation due to their expressive power and generation accuracy [8]. The realization of texture representation using image generation methods based on deep learning will make it possible to reproduce and edit a wide variety of textures.

In this study, we propose a BTF prediction model using a Deep Neural Networks (DNN) as a first attempt to generate textures based on affective texture perception. We also verify the accuracy and clarify the effectiveness of the proposed method.

David C. Wyld et al. (Eds): CMC, SOFT, MLT, SAI - 2022
pp. 45-53, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.120505

2. RELATED WORK

In the field of computer graphics (CG), techniques for reproducing realistic images have been developed. Since accurately reproducing the appearance of an object's surface is essential for realistic computer graphics, BRDF and BTF, which represent the characteristics of reflections on an object's surface, are used as reproduction techniques. Especially for BTF-based systems, a large number of texture images are collected under diverse conditions using specialized acquisition equipment in advance [4, 9]. However, time and huge amounts of data are problems for collection by measurement devices. To solve this problem, studies on BTF generation using DNNs are being reported [8, 5]. since BTFs require a large amount of data to be measured in advance, it is very useful to reproduce the entire BTF data set from a small number of BTFs.

On the other hand, researchers have conducted interdisciplinary texture studies in psychophysics, engineering, and brain physiology [10,11]. Some examples of research on rendering methods include the use of BTF for cloth [12,13]. However, in these studies, BTF is measured using measurement devices, and it takes a great deal of time to perform the measurement. In this study, we attempt to reduce the measurement time by predicting and interpolating the BTF.

In addition, image generation has been an active area of research related to deep learning, and the latest studies have shown remarkable results [14]. In recent years, research on texture generation using GANs has been reported [15], but no research using time series data has been reported. However, it has been reported that Predictive-coding based DNN (PredNet [16]) (Fig.1), a learning model for time series data, can predict optical illusions, which is one of the perceptual processes of the brain [17]. Therefore, there is a possibility that we can predict texture, which is a cognitive processing of the brain, by applying this model. In this study, we attempt to predict the change in texture due to changes in the angle of view using the PredNet framework. In this study, we use the BTF dataset, in which the viewpoint angle changes continuously, as the training data.

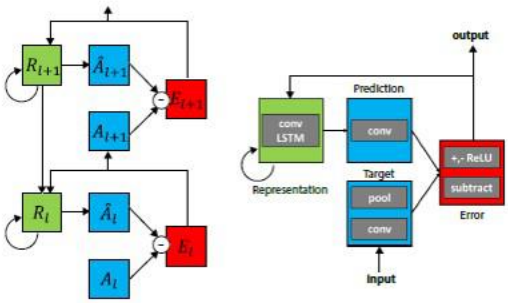


Figure 1. PredNet [16]

3. METHOD

In this study, we propose a BTF prediction model using PredNet. Fig. 2 shows an overview of the proposed method. The PredNet used in this method is a DNN with time-series image data as input and output. The method uses a series of continuously varying viewpoint angles of a texture image as the input signal. This method enables the prediction of texture images with continuously changing angles.

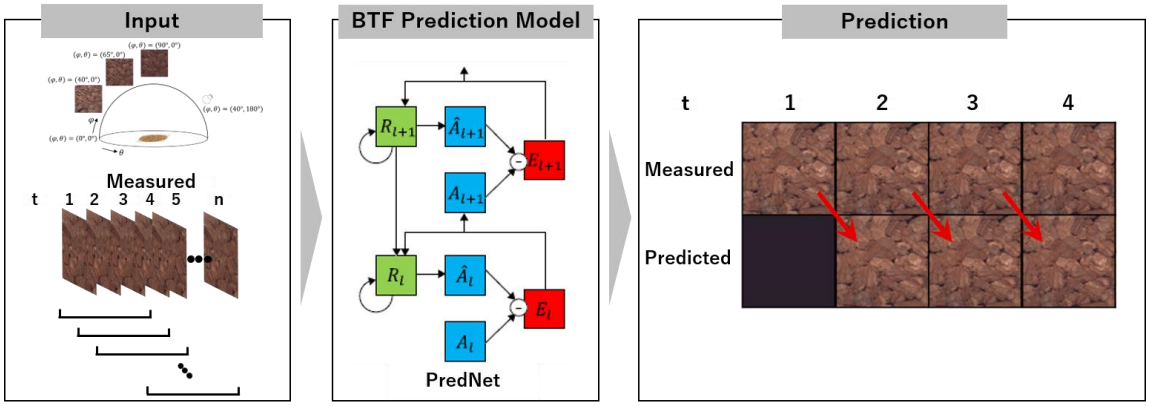


Figure 2. Concept of our proposal method

3.1. Material

The target material was measured using an OGM-CCD camera measuring device manufactured by Digital Fashion. The OGM-CCD consists of a sample stand, a metal halide light source, and a CCD camera, and it is capable of measuring the optical axis in two axes and the sample stand in two axes, for a total of four axes of angular rotation. In this study, we use the images this device captures as BTF. Table 2 shows this system's measurement conditions. We measured the

Table 1. Measurement conditions

Viewing		Illumination	
Elevation	Azimuth	Elevation	Azimuth
30~90° (every 1°)	0, 90, 180, 270°	30~90° (every 5°)	0, 90, 180, 270°

Table 2. Measurement materials

Conditions			Materials		
	Reflection property	Pattern	Textile	Wood	Paper
A	Diffuse	Irregular			
B	Diffuse	Regular			
C	Specular	Irregular			
D	Specular	Regular			

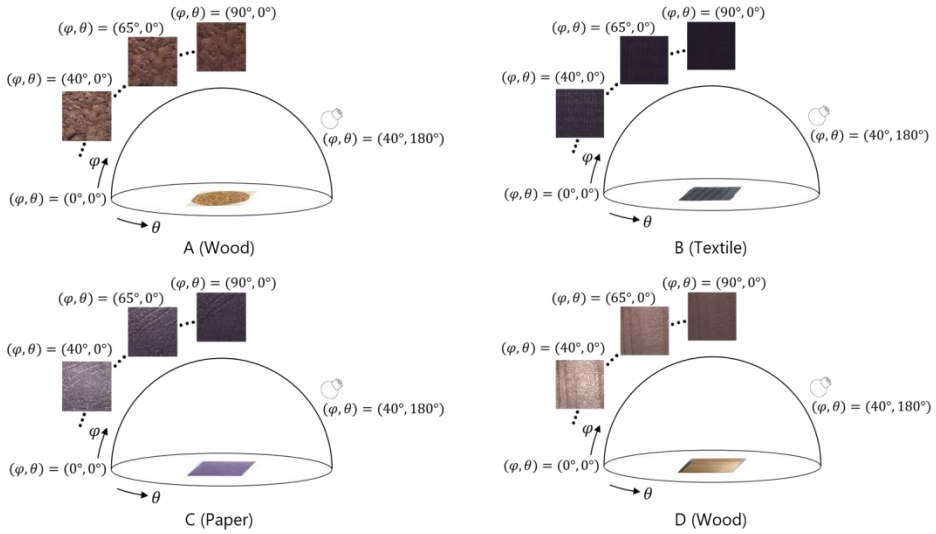


Figure 3. Measurement method and obtained images

elevation angle of the viewpoint at every 1◦ , the azimuth angle between the viewpoint and the light source at every 90◦ , and the elevation angle of the light source at every 5◦.

We selected textile, wood, and paper as the measurement materials for BTF. For each material, we treated four conditions: material with diffuse reflection property and irregular pattern (condition A), material with diffuse reflection property and regular pattern (condition B), material with specular reflection property and irregular pattern (condition C), and material with specular reflection property and regular pattern (condition D). We measured a total of 12 materials. Table 1 shows the materials we used for the measurements.

We obtained Each of the 11,396 BTFs for each material. Fig.3 shows some of the measurement results for each material. Fig.3 shows the visibility of the viewpoint in relation to the position of the viewpoint and the light source.

3.2. Dataset

Based on the measured BTFs, we created a dataset for learning, verification, and testing. We sampled each material at 2◦ intervals so the change in viewpoint elevation angle would be continuous, and we used 10 images as one set (Fig.4). In this study, we used 66336 sets for training, 8292 sets for verification, and 691 sets to test each material. We divided the data sets randomly.

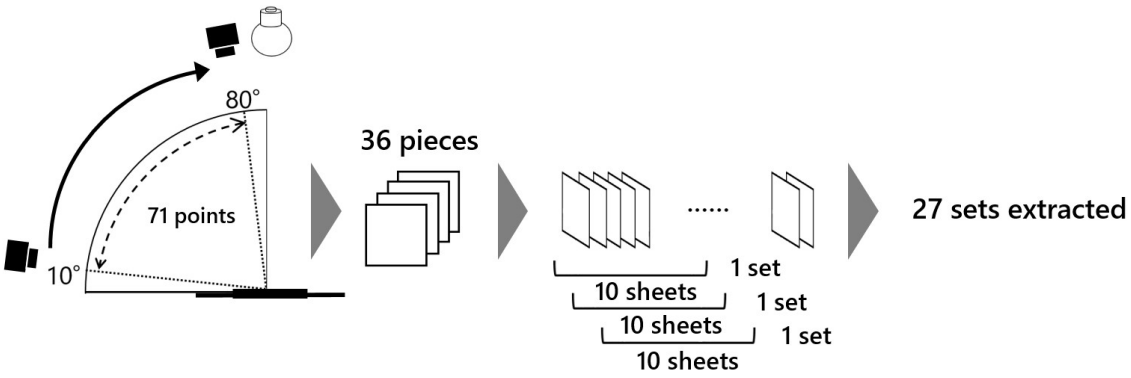


Figure 4. Overview of dataset creation

4. RESULT AND DISCUSSION

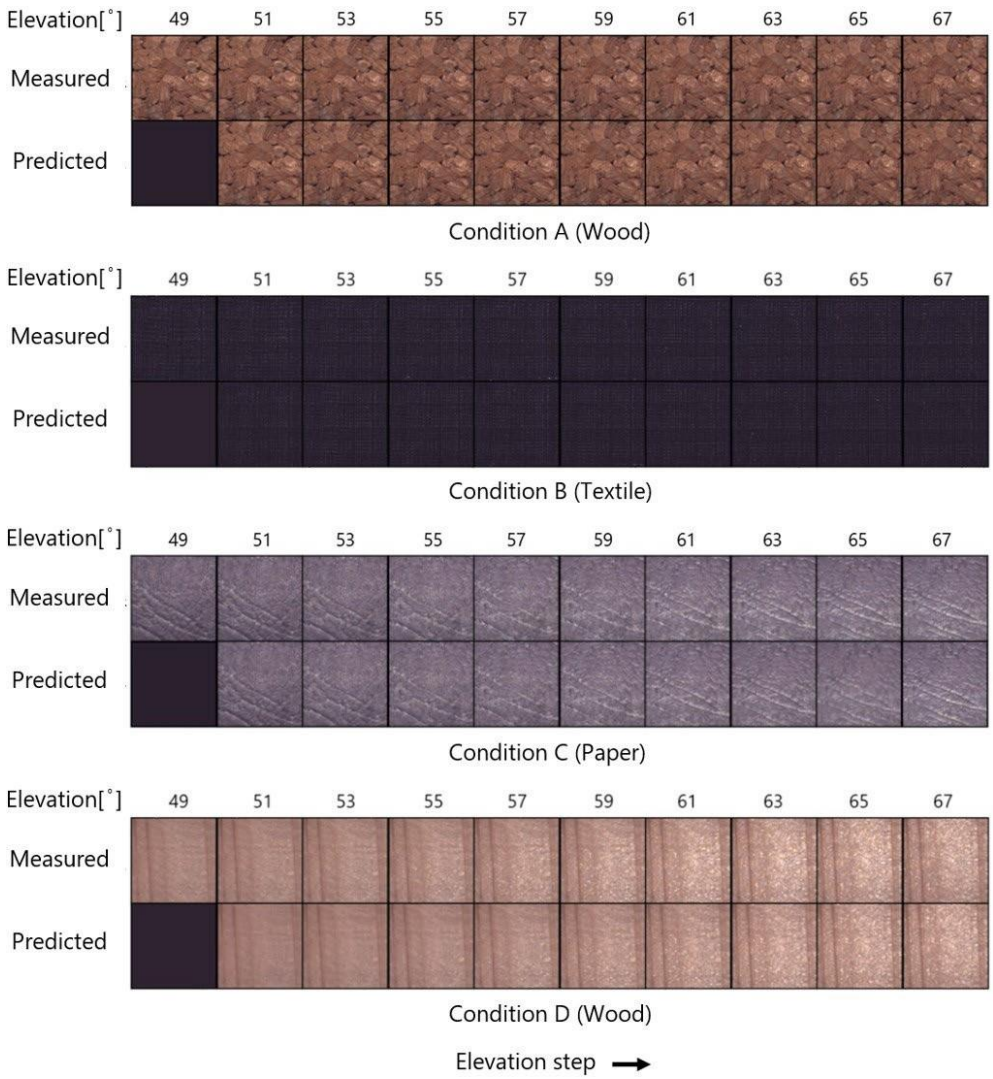


Figure 5. Example of prediction results (Azimuth of viewpoint: 180◦, Elevation of light source: 65◦, Azimuth of light source: 0◦)

We conducted the prediction of future frames using the series of images as input data. We used the BTFs measured in the previous section as the input images to the network. The network structure has four layers, a convolutional filter size of 3 × 3, and the number of filters is 3, 48, 96, and 192, starting from the lower layers. Adam was used as the model optimization algorithm. The hyperparameters for learning were set based on previous studies [16].

Examples of BTF prediction results are shown in Fig.5. For all materials and conditions, the predicted images are similar to the measured images. We confirmed that predictive capabilities of PredNet are effective in generation BTFs.

Furthermore, to quantitatively verify the effectiveness of the proposed method, we evaluated the accuracy of BTF prediction for each material. We used SSIM [18] to evaluate accuracy; SSIM is an index that assumes that the similarity of image structure contributes to near human image quality degradation; the similarity is calculated from luminance, contrast, and structure. Image

Table 3. Average of SSIM in test data

Conditions			Materials			Average of conditions
	Reflection	Pattern	Textile	Wood	Paper	
A	Diffuse	Irregular	0.982	0.987	0.991	0.987
B	Diffuse	Regular	0.986	0.986	0.984	0.986
C	Specular	Irregular	0.923	0.982	0.990	0.965
D	Specular	Regular	0.947	0.971	0.985	0.968
Average of materials			0.960	0.982	0.988	

compression guidelines for digitized documents state that SSIM is indistinguishable from the original image at 0.98 or higher, degradation is visible when enlarged at 0.90-0.98, and degradation is obvious at 0.90 or lower [19]. Table 3 shows the average SSIM values in the test data. The table shows that the values are high under all conditions. The SSIM results thus confirm that the reflective characteristics and patterns are accurately predicted under all conditions.

A comparison of materials shows that the values for cloth are lower than those for wood and paper. When we look at the SSIM of each of the 12 materials, we find that the SSIM of fabrics with specular reflectance characteristics is particularly low. A comparison of the different conditions shows that the condition with specular reflectance characteristics has a slightly lower SSIM value than the condition with diffuse reflectance characteristics. To examine these factors, we focused on the SSIM for each angle of the predicted results (Fig.6). Fig.6 shows that the SSIM values are high for each viewpoint elevation angle for conditions A and B, which are diffuse reflection characteristics. On the other hand, the SSIM values are low in the specular reflection conditions, condition C and D, where specular reflection begins to appear. Comparing the measured and predicted images for conditions C and D in Fig.7 with a viewpoint elevation angle of 55◦ , where the change in SSIM values is large, we confirmed that the predictions of the glossy areas were not sufficient. These results suggest that, due to the characteristics of PredNet, this method is not accurate enough to predict steep pixel value changes that would cause specular reflections to appear. Therefore, a new model that shows high accuracy even in steep changes due to angular changes will be considered.

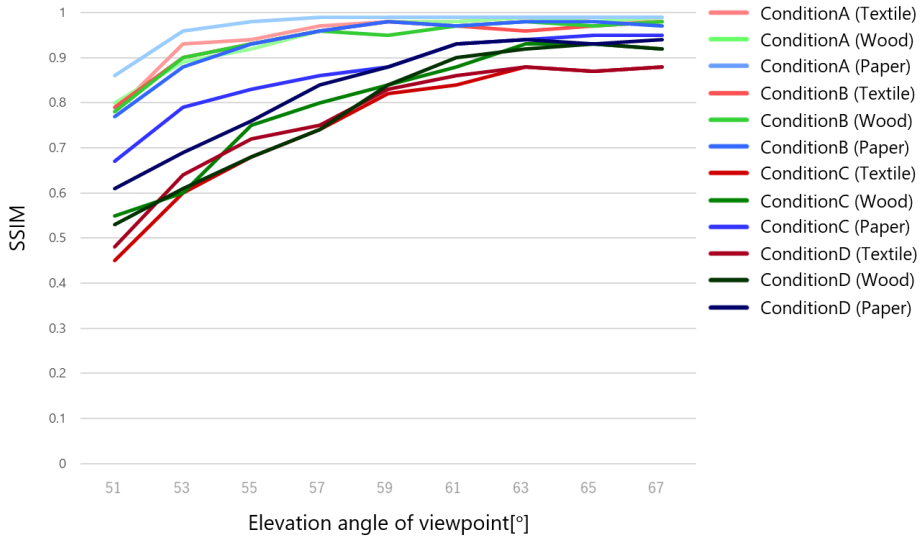


Figure 6. SSIM with change in elevation angle of viewpoint

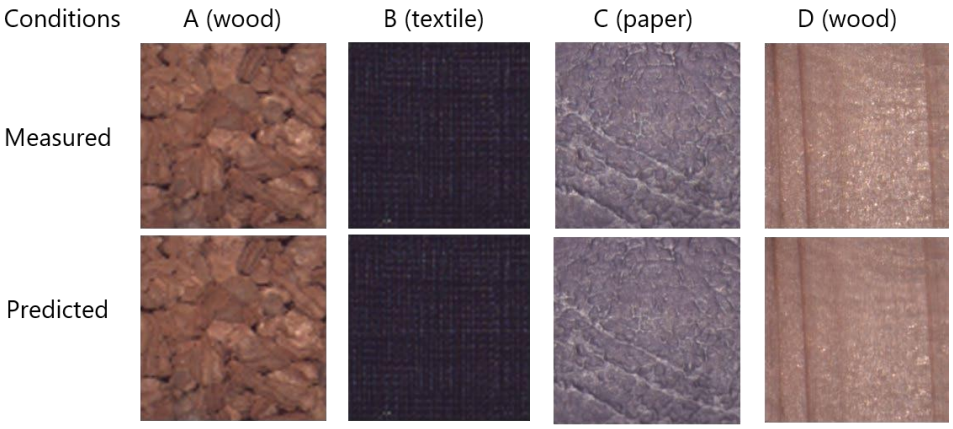


Figure 7. Measured and predicted images for a viewpoint elevation angle of 55◦

5. CONCLUSION AND FUTURE WORK

In this study, we developed a BTF prediction model for texture generation based on affective texture perception. We used PredNet for prediction, and build the model by training a BTF dataset with a series of continuously varying angles. We determined the proposed method's effectiveness by calculating the predicted images' SSIM for 12 materials with various optical properties. The results show that the proposed method is effective for predicting diffuse reflection optical properties and irregular and regular patterns.

In the future, we will study a new model that shows high prediction accuracy even for steep texture changes due to angular changes.

ACKNOWLEDGEMENTS

This work was supported in part by the Center of Innovation Program from Japan Science and Technology Agency, JST.

REFERENCES

[1] N. Sunda, K. Tobitani, I. Tani, Y. Tani, N. Nagata, N. Morita (2020) Impression estimation model for clothing patterns using neural style features. HCI International 2020 - Posters (HCII2020), CCIS 1226, 689-697.

[2] K. Tobitani, T. Matsumoto, Y. Tani, H. Fujii, N. Nagata, Modeling of the Relation between Impression and Physical Characteristics on Representation of Skin Surface Quality, The Journal of The Institute of Image Information and Television Engineers, 71, 11, 2017

[3] F. Nicodemus, J.Richmond, J. Hsia, I. Ginsberg, T. Limperis: ”Geometrical considerations and nomenclature for reflectance”, NBS monograph, 160, (1977)

[4] K. J. Dana, B. Van Ginneken, S. K. Nayar, J. J. Koenderink: “Reflectance and texture of real- world surfaces”, ACM Transactions On Graphics (TOG), 18, 1, pp.1-34 (1999)

[5] G. Rainer, W. Jakob, A. Ghosh, T. Weyrich: “Neural btf compression and interpolation”, In Computer Graphics Forum, 38, 2, pp.235-244 (May, 2019)

[6] I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al.: ” Generative adversarial nets”, Advances in neural information processing systems, 27 (2014)

[7] D. P. Kingma, M. Welling: “Auto-encoding variational bayes”, arxiv preprint arXiv:1312.6114 (2013)

[8] X. Zhang, J. Dong, Y. Gan, H. Yu, L. Qi: “BTF data Generation based on Deep Learning”, Procedia computer science, 147, pp.233-239 (2019)

[9] J. Y. Han and K. Perlin: “Measuring bidirectional texture reflectance with a kaleidoscope”, in ACM SIGGRAPH 2003 Papers, pp.741-748 (2003)

[10] I. Motoyoshi, S. Nishida, L. Sharan, and E. H. Adelson: “Image statistics and the perception of surface qualities”, Nature, 7174, pp.206-209 (2007)

[11] S. Nishida: “New Academic Field Research (Inovative SHITSUKAN science and technology)”, Basic Psychological Research, 37, 1, pp.117-118 (2018)

[12] Y. Takeda， H. Tanaka : “Reflective Modeling of Gold Threads in Digital Content Creation of Noh Costume”， The collection of contributions to the annual convention of the Society of Portrait Electronics, 35, pp.197-202（2007）, Japanese

[13] Y. Takeda， H. Tanaka ： “Multi-resolution anisotropic BTF modeling of gold brocade using multi-directional illumination HDR images.” ， Transactions of the Institute of Electronics,Information and Communication Engineers D, 91,12, pp.2729-2738（2008）, Japanese

[14] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, T. Aila: “Analyzing and improving theimage quality of stylegan”, In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.8110-8119 (2020)

[15] Y. Guo, C. Smith, M. Hasan, K. Sunkavalli, S. Zhao: “MaterialGAN: reflectance capture using a generative SVBRDF model”, arXiv preprint arXiv:2010.00114 (2020)

[16] W. Lotter, G. kreiman, D. Cox: “Deep predictive coding networks for video prediction and unsupervised learning”, arXiv preprint arXiv:1605.08104 (2016)

[17] E. Watanabe, A. Kitaoka, K. Sakamoto, M. Yasugi, K. Tanaka: “Illusory motion reproduced by deep neural networks trained for prediction”, Frontiers in psychology, 9, 345 (2018)

[18] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli: ”Image quality assessment: from error visibility to structural similarity”, IEEE transactions on image processing, 13, 4, pp.600-612 (2004)

[19] M. Kobako: “Road to Standardization Standardization Committee Report (3) Image Compression Guidelines for Electronic Documents”, 50, 5, pp.21-24, (2011)

AUTHORS

SOICHIRO KIMURA

Soichiro Kimura received the BR degree in human system interaction from Kwansei Gakuin University in 2020. He is currently a second year MS student in Science and Technology cources, Kwansei Gakuin University. His research interests include Kansei (affective) information and machine learning.

KENSUKE TOBITANI

Kensuke Tobitani has been working as an Associate Professor in the Department of Information Systems, University of Nagasaki, Nagasaki, JAPAN. of interest includes machine learning, affective engineering, and computer vision. The author's main research theme is the analysis of human sensibility by machine learning, and he has been engaged in research on modeling the relationship between physical properties of objects and the human In particular, he has published papers on understanding and analyzing the texture of objects and estimating human impressions of three-dimensional objects.

NORIKO NAGATA

Noriko Nagata received a BS degree in mathematics from Kyoto university in 1983 and a PhD degree in systems engineering from Osaka University in 1996. She was a researcher at the Industrial Electronics and Systems Laboratory of Mitsubishi Electronics Corporation from 1983 to 2003. She joined Kwansei Gakuin University in 2003 as an associate professor. She is currently a professor in the department of human system interaction and a director of the reseach center for Kansei Value Creation. In 2009, she was a visiting scholar at Purdue University. Her research interests include Kansei (affective) information processing, computer graphics, and multimedia systems. She is a member of the IEEE and ACM.

AUTHOR INDEX