Group01

Pic2Model Vision Document

Version <1.0>

Pic2Model	Version: <1.0>
Vision (Small Project)	Date: 07/11/2024
<document identifier=""></document>	

Revision History

Date	Version	Description	Author
07/11/2024	1.0	This version includes the initial overview of the project's vision, outlining the purpose, scope, objectives, and key requirements.	All team's members

Pic2Model	Version: <1.0>
Vision (Small Project)	Date: 07/11/2024
<document identifier=""></document>	

Table of Contents

1.	Introduction		4
2.	Posit	tioning	5
	2.1	Problem Statement	5
	2.2	Product Position Statement	5
3.	Stake	eholder and User Descriptions	6
	3.1	Stakeholder Summary	6
	3.2	User Summary	6
	3.3	User Environment	7
	3.4	Alternatives and Competition	7
4.	Prod	luct Features	9
5.	Non-	-Functional Requirements	10
	5 1	Detailed Requirements	10

Pic2Model	Version: <1.0>
Vision (Small Project)	Date: 07/11/2024
<document identifier=""></document>	

Vision (Small Project)

1. Introduction

This section provides a comprehensive overview of the project, detailing its purpose, scope, stakeholders, and key terms used throughout the document.

Purpose: The Pic2Model project aims to address the needs of various users across education, e-commerce, tourism, and design sectors. The reasons for undertaking this project include:

- Enhancing Educational Quality: Providing interactive 3D models to help teachers and students understand complex concepts more effectively.
- **Improving Online Shopping Experience**: Allowing customers to view products from all angles before making a purchase.
- **Creating Virtual Travel Experiences**: Enabling people to explore destinations without the need to travel physically.
- **Supporting Design Work**: Assisting designers in creating accurate and detailed 3D models quickly.

End Users:

- Teachers and students
- Online shoppers
- Virtual tourists
- Designers and artists

Target User Information: End users need intuitive and effective visual tools to enhance their learning, shopping, and design experiences.

Key Functional and Non-Functional Requirements:

- **Functional Requirements**: Create 3D models from 2D images and text descriptions, provide interactive 3D experiences, offer AI-powered design tools.
- Non-Functional Requirements: Ensure security, high processing speed, scalability, and user-friendliness.

Scope: The scope of this Vision Document (PVD) includes describing the system's functions and requirements. It details end-user needs, core system functionalities, and non-functional requirements necessary for the system to operate effectively.

References:

- <u>Mapillary. (n.d.). GitHub mapillary/OpenSfM: Open source Structure-from-Motion pipeline.</u> GitHub.
- Autonomousvision. (n.d.). GitHub autonomousvision/differentiable volumetric rendering: This
 repository contains the code for the CVPR 2020 paper "Differentiable Volumetric Rendering:
 Learning Implicit 3D Representations without 3D Supervision." GitHub.
- Amnah's Lab. (2023, October 20). Gentle Introduction to Point Cloud Registration using Open3D [Video]. YouTube.
- Nicolai Nielsen. (2022, February 22). Elevate Your Point Cloud Game with State-of-The-Art Depth Neural Networks in Open3D and OpenCV [Video]. YouTube.
- Project Vision Document. (n.d.). BAC TRAINING & CONSULTANCY VN.

Pic2Model	Version: <1.0>
Vision (Small Project)	Date: 07/11/2024
<document identifier=""></document>	

- LN04 Project Assignments (n.d). Moodle
- Blog "Alpha3D Reviews"
- Blog "CSM Reviews"
- Blog "Kaedim Reviews"

2. Positioning

2.1 Problem Statement

The problem of	lack of virtual tools and educational aids to support citizens in various aspects of life, especially in work and entertainment
affects	people's ability to have realistic experiences and understand real-world problems or objects
the impact of which is	leads to misguided decisions, wasting time, effort, and money, and creating unsatisfactory experiences for users
a successful solution would be	simulating realistic models to assist citizens in work, study, and entertainment
	enhancing realistic experiences for citizens by allowing them to interact with virtual models created by the website
	saving time, effort, and money while still providing satisfactory experiences and suitable products for their needs

2.2 Product Position Statement

For	Almost citizens, especially,	
	rimost citizons, especially,	
	 Teachers, students 	
	 Travelers 	
	 Designers 	
Who	need virtual tools or 3D – designs for studying, working, entertainment,	
The Pic2Model	is a web app	
That	converts 2D images into detailed 3D models or create a detailed 3D models from a description text	
	helps users provide interactive and customizable 3D content	
Unlike	Alpha3D, CSM AI, Image to STL, Kaedim	
Our product	offers accessibility, AI integration, and versatile output formats, making 3D modeling more user-friendly and adaptable to various needs.	

Pic2Model	Version: <1.0>
Vision (Small Project)	Date: 07/11/2024
<document identifier=""></document>	

3. Stakeholder and User Descriptions

3.1 Stakeholder Summary

Name	Description	Responsibilities	
Development Team	A group of students directly involved in the software development process, responsible for all stages from requirement analysis to deployment. - Conducts requirement and design, programming, testing deployment - Ensures that development requirements and goals - Collaborates on troublesh enhancements, and timely oproject phase.		
Theory Lecturer (Trần Duy Hoàng) The lecturer is responsible for supervising and mentoring the team in the software development process. Provides foundational knowledge and practical insights.		- Guides students on industry best practices and theoretical foundations in software development - Acts as a supervisor to ensure adherence to software development standards and methodologies - Assesses students' understanding and application of theoretical knowledge.	
Teaching Assistant (Ngô Ngọc Đăng Khoa)	A TA who communicates regularly with students, providing timely answers to questions and offering advice throughout the development process.	 Acts as a primary point of contact for student queries and feedback Provides ongoing guidance to help students meet assignment standards Reviews student work and offers constructive suggestions prior to final submission on Moodle. 	

3.2 User Summary

Name	Description	Responsibilities	Stakeholder
Students & Educators	In educational settings, especially those centered around visual learning and creative exploration, being able to visualize objects in multiple angles is crucial. Whether it's for science, engineering or art.	 - Upload / Capture image(s) of the object they want to generate. - Create visual presentations or reports. 	Designer Implementer
3D Designers	With the rapid advancement of digital world and users' demand, 3D graphic designers need to	- Upload / Capture image(s) of the object they want to generate.- Create visual presentations or reports.	Designer Implementer

Pic2Model	Version: <1.0>
Vision (Small Project)	Date: 07/11/2024
<document identifier=""></document>	

	constantly design 3D objects for many use cases, some of those replicate real-world objects.		
Tourists	When tourists want to view historical buildings. historical sites as the whole, they need something to generate 3D view of these from image(s)	 - Upload / Capture image(s) of the object they want to generate. - Create visual presentations or reports. 	Designer Implementer
Shoppers	When shoppers buy something, they want to try it on or place it somewhere to consider whether to buy it or not. They need the 3D view of that object	 - Upload / Capture image(s) of the object they want to generate. - Create visual presentations or reports. 	Designer Implementer

3.3 User Environment

Working environment	Description
Number of people involved in completing the task	- Just 1 person can complete the task - This will not change.
Task cycle	- Spend less than 1 minute for each activity - If the object want to generate 3D view too big, users need to wait several minutes.
Unique environmental constraints	- No constraint
Platform	- Any devices have access to the Internet can use - In the future, attempting to give access to off-line users

3.4 Alternatives and Competition

In the market for transforming 2D images into 3D models, there are several alternative solutions and competitors available to stakeholders, each with unique strengths and weaknesses. These include commercial products like Alpha3D, CSM AI, and Kaedim. Below is a summary of each competitor's perceived capabilities and limitations as observed by stakeholders.

(1) Alpha3D

Alpha3D is a revolutionary generative AI platform tailored for transforming text prompts and 2D images into detailed 3D digital assets. This innovative tool is specifically designed to cater to the needs of game developers, content creators, and hobbyists who may not have extensive skills in 3D modeling. By leveraging cutting-edge AI technology, Alpha3D enables users to produce high-quality, game-ready 3D models swiftly and efficiently.

• Pros:

Pic2Model	Version: <1.0>	
Vision (Small Project)	Date: 07/11/2024	
<document identifier=""></document>		

- + **Speed of Production**: Accelerates the creation process, enabling users to go to market faster with their 3D content.
- + Accessibility: Opens the field of 3D content creation to those without specialist knowledge or skills.
- + Cost Savings: Offers a more budget-friendly solution compared to conventional 3D modeling techniques.
- + **Initial Free Assets**: New users can generate their first 50 3D assets for free, providing a risk-free opportunity to assess the tool's capabilities.

Cons:

- Limited Object Categories: The AI model is currently trained primarily in two categories: shoes and furniture, which may limit versatility.
- Dependence on Input Quality: The quality of the generated 3D model may be heavily reliant on the quality
 of the input image or text description.
- Potential for Inaccuracy: While AI technology has made strides, there may be occasional inaccuracies in the translation from 2D to 3D.

(2) **CSM AI**

CSM.ai, also known as **Common Sense Machines (CSM)**, is an advanced AI product that allows users to generate controllable and game-engine ready 3D animated worlds from single images, text, and sketches. It's designed to turbocharge 3D workflows and can generate 3D worlds from multimodal inputs including images, sketches, and text.

Pros:

- + **Innovative 3D Generation**: The ability to generate 3D worlds from text, images, or sketches is unparalleled, offering a new dimension of creativity.
- + **Flexible and User-Friendly**: The multimodal input system and intuitive interface cater to both beginners and professionals alike.
- + **Rapid Prototyping**: Significantly reduces the cycle time from concept to 3D print or game prototype, enabling faster testing and development.
- + Community and Customization: Encourages a collaborative environment with the option for users to choose or create unique styles for their 3D assets.

Cons:

- Learning Curve: New users may require time to fully grasp the capabilities and make the most of the advanced features offered.
- Resource Intensity: High-quality 3D rendering may demand significant computational resources, potentially limiting access for users with less powerful hardware.
- Pricing Clarity: For those seeking to integrate CSM.ai into large-scale projects, the lack of transparent pricing for studio and enterprise solutions may necessitate direct inquiry.
- Limitations in Free Version: The free version of CSM.ai has certain limitations.

(3) Kaedim

Pic2Model	Version: <1.0>
Vision (Small Project)	Date: 07/11/2024
<document identifier=""></document>	

Kaedim is an innovative AI-powered tool that transforms 2D images into high-quality 3D models in a matter of minutes. It's designed to streamline the 3D modeling process, making it accessible to a wider range of users, including game developers, graphic designers, and digital artists.

Pros:

- + Rapid 3D Model Creation: Quickly generate detailed 3D models from simple 2D images.
- + **Automatic Texturing**: Automatically applies textures to models for a realistic appearance.
- + **Production-Ready Models**: Produces high-quality models that are ready for use in games and other 3D applications.
- + **Seamless Integration**: Integrates with popular 3D modeling software for added flexibility.

• Cons:

- Limited Accuracy: May not always accurately capture the original shape or style of the 2D image.
- **Difficulty with Complex Scenes**: Might struggle with complex or dynamic scenes.
- Subscription-Based Model: Requires a subscription or trial to access all features.
- Limited Customization: Offers limited customization and editing options.

4. Product Features

ID	Products Features	Description	Priority
1	Photo Upload & Capture	Allows users to upload multiple photos or take real-time snapshots via a webcam or mobile camera. The app provides guidelines to help users capture photos from optimal angles for accurate 3D modelling.	High
2	3D Model Generation	Process the uploaded images to generate a high-resolution 3D model with accurate representation of features. This feature uses AI-driven algorithms to ensure model realism and detail fidelity.	High
3	Interactive Preview	Provides a preview window where users can rotate, zoom, and pan around the generated 3D model, allowing users to inspect model accuracy and adjust as needed before finalizing.	High
4	Model Customization	Offers tools for users to make basic adjustments to the 3D model, such as rotation, scaling, and adding accessories or textures, enhancing personalization and usability for various applications.	Medium
5	Export & Download Options	Enables users to download the 3D models in various file formats (. OBJ, STL) compatible with other design and animation tools, allowing for easy integration into external workflows.	High
6	User Account Management	Provides secure account functionality where users can register, log in, and manage their personal information and previously created 3D models, facilitating easy access and retrieval of past projects.	Medium

Pic2Model	Version: <1.0>
Vision (Small Project)	Date: 07/11/2024
<document identifier=""></document>	

7	Storage for Model History	Saves user-generated 3D models, enabling users to access and modify past models without re-uploading photos, improving efficiency for repeat use or iterative design work.	Medium
8	Real-Time Processing Feedback	Offers users a status indicator or progress bar during model generation, managing user expectations on processing time and reducing the likelihood of abandoning the process.	Low
9	Cross-Platform Compatibility	Ensures that the web app is accessible on both desktop and mobile browsers, with responsive design and support for major browsers (Chrome, Firefox, Safari), reaching a wider user base.	High
10	Guided Model Capture Tutorial	Includes a brief tutorial or tips on how to take optimal photos for best 3D modelling results, improving the accuracy and quality of the final model through user education.	Low

5. Non-Functional Requirements

To ensure a seamless user experience and robust platform, we've identified the following key -non-functional requirements:

- **Scalability**: The platform must handle concurrent requests from multiple users without performance degradation, especially during peak times.
- **Security and Privacy**: Ensure data security, with robust encryption for uploaded images, 3D models, and user information, adhering to data protection regulations.
- **Performance**: Aim for rapid model generation times, with minimal lag during image processing and model preview interactions.
- **Usability**: Provide an intuitive interface, with clear instructions and minimal steps to guide users through the upload, model generation, and download processes.
- Compatibility: Support modern web browsers (Chrome, Firefox, Safari) and devices (desktop and mobile) to reach a broad audience.
- Availability and Reliability: The system should have minimal downtime and backup options to prevent data loss in case of unexpected outages.

5.1 Detailed Requirements

1. Scalability

- **Description**: The system must handle increased user load without significant performance drops. As the user base grows, the infrastructure should support additional concurrent requests and maintain processing speeds.
- **Implementation**: For now, we are considering improving our own database before looking for more advanced options like cloud-based

2. Security and Privacy

- **Description**: Ensuring user data protection is paramount, particularly for uploaded photos and 3D models. This includes robust encryption methods for data at rest and in transit.
- Implementation:
 - Use **HTTPS** for secure communication.
 - o Implement strong authentication for user accounts.

Pic2Model	Version: <1.0>	
Vision (Small Project)	Date: 07/11/2024	
<document identifier=""></document>		

- o Encrypt uploaded images and models using advanced algorithms (e.g., AES-256).
- Comply with data protection regulations such as GDPR to ensure user privacy rights are maintained.

3. Performance

- **Description**: The app should process and render 3D models efficiently, with minimal wait times for users. Target maximum model generation time is under **60 seconds** for standard photo sets.
- **Implementation**: Optimize algorithms for image processing and 3D reconstruction. Employ techniques like **parallel processing** and GPU acceleration to speed up heavy computations.
- Performance Monitoring: Implement real-time performance tracking and alerts to identify and mitigate slowdowns.

4. Usability

- **Description**: The app interface should be intuitive and user-friendly, enabling users with minimal technical expertise to navigate and use the system effectively.
- Implementation:
 - Use consistent UI/UX patterns.
 - o Provide clear instructions and tooltips.
 - Ensure accessibility features are integrated (e.g., compatibility with screen readers, keyboard navigation).
- **Usability Testing**: Conduct user feedback sessions and A/B testing to refine the interface based on user preferences and ease of use.

5. Compatibility

- **Description**: The web app must function seamlessly across different devices and browsers to cater to a wider user base.
- Implementation:
 - Ensure cross-browser compatibility by testing on Chrome, Firefox, Safari, and Edge.
 - Design with responsive frameworks (e.g., Bootstrap) to adapt to various screen sizes, including desktops, tablets, and smartphones.

6. Availability and Reliability

- **Description**: The platform should provide a high level of availability, with minimal downtime to ensure continuous access for users. The target uptime is **99.9%**.
- Implementation:
 - Use redundancy in cloud services and databases.
 - o Implement a disaster recovery plan with automated data backups.
- Reliability Monitoring: Implement tools like New Relic or Datadog for monitoring uptime and service health.

7. Maintainability

- **Description**: The system should be easy to update and maintain, allowing developers to make improvements or fix issues with minimal disruption.
- Implementation:
 - Use modular code structures and follow coding best practices.
 - Implement automated testing and continuous integration (CI/CD) pipelines for streamlined updates.
- **Documentation**: Maintain comprehensive documentation for developers and stakeholders.

8. Scalability for Future Growth

Pic2Model	Version: <1.0>	
Vision (Small Project)	Date: 07/11/2024	
<document identifier=""></document>		

- **Description**: The architecture should support additional features or enhancements in the future without significant rework.
- Implementation:
 - o Use microservices or service-oriented architecture (SOA) to modularize services.
 - o Plan for API integrations that allow easy expansion of capabilities.