Group01

Pic2Model
Vision Document

Version 1.1

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
14/11/2024	1.1	Detail the product's functional and non-functional requirements, alternatives and competitors, and the user environment based on TA’s comments in PA1.	Vũ Minh Phát, Vũ Mạnh Quân

Table of Contents

Vision (Small Project)

1. Introduction

1.1 Purpose

The purpose of this document is to collect, analyze, and define high-level needs and features of the **Pic2Model**. It focuses on the capabilities needed by the stakeholders and the target users, and **why** these needs exist. The details of how the **Pic2Model** fulfills these needs are detailed in the use-case and supplementary specifications.

The introduction of the **Vision** document provides an overview of the entire document. It includes the purpose and references of this **Vision** document.

1.2 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.

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.

1.3 References

Applicable references are:

- Mapillary. (n.d.). GitHub - mapillary/OpenSfM: Open source Structure-from-Motion pipeline. 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.
- Slide PowerPoint Introduction to Software Engineering
- Blog "Alpha3D Reviews"
- Blog "CSM Reviews"
- Blog "Kaedim Reviews"

2. Positioning

2.1 Problem Statement

The problem of	Lack of virtual tools and educational aid 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 all citizens, especially, • Teachers, students • Travelers • Designers
Who	Need virtual tools or 3D – designs for studying, working, entertainment, etc.
The Pic2Model	is a web app
That	Converts 2D images into detailed 3D models or create a detailed 3D model 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.

3. Stakeholder and User Descriptions

3.1 Stakeholder Summary

Name	Description	Responsibilities
Development Team	All members of the group.	Develop front-end and back-end for the product. Implement new features. Maintain existing functionality. Optimize platform performance. Fix bugs and technical issues. Integrate user feedback into the development process.
Product Manager	A member in the group.	Manage the team, project process. Ensure the website meets business goals.
Business Stakeholders	Investors and strategic partners.	Provide funding and resources. Set strategic direction. Monitor platform growth and metrics. Approve major feature developments. Evaluate business opportunities. Review financial performance.

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 constantly design 3D objects for many use cases, some of those replicate real-world objects.	Upload / Capture image(s) of the object they want to generate. Create visual presentations or reports.	Designer Implementer
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:**
  - + **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

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	Product Features	Description	Priority
1	User Registration & Authentication	Allow users to register, log in, and manage their accounts securely. Includes email verification, password reset, and account blocking for suspicious activities.	High
2	Photo Upload & Capture	Enables users to upload multiple photos from different angles or capture real-time snapshots via webcam or mobile. Offers guidelines for optimal 3D modeling.	High
3	3D Model Generation	Uses AI algorithms to generate high-resolution 3D models from uploaded images, ensuring detailed and accurate results.	High
4	Interactive 3D Model Preview	Users can interact with the generated 3D models through rotation, zooming, and panning for a 360-degree inspection before finalizing.	High
5	Cross-Platform Compatibility	Ensures the web app works seamlessly on both desktop and mobile browsers, supporting Chrome, Firefox, Safari, and responsive designs.	High
6	Export & Download Options	Allows users to export 3D models in various formats (OBJ, STL, PLY) or as compressed files (zip). Includes options for exporting images and videos.	High
7	Subscription & Premium Upgrade	Users can upgrade to premium accounts for additional features and benefits. Supports payment options and promotional gift codes.	High
8	User Profile Management	Users can view and edit personal information (except email and password). Provides options for account deletion with a 72-hour grace period.	Medium
9	Premium Photo Upload	Premium users can upload a single high-quality image for faster 3D model generation, improving efficiency for quick designs.	Medium
10	Model Customization Tools	Offers basic tools for modifying 3D models, including rotation, scaling, and adding accessories or textures for enhanced personalization.	Medium

11	Prompt-Based 3D Model Generation	Premium users can write prompts to generate customized 3D models, supporting both positive and negative prompts for detailed customization.	Medium
12	Model History & Storage	Saves user-generated 3D models for easy retrieval and modification, allowing users to revisit past projects without re-uploading images.	Medium
13	Model Sharing & Collaboration	Provides a workshop page where users can share and showcase their created 3D models. Enable users to explore others' work for inspiration.	Medium
14	Notification & Email Alerts	Send email notifications for account registration, premium subscription upgrades, cancellations, and other account activities.	Medium
15	Gift Code Purchase & Management	Premium users can buy gift codes for others, enabling gift code management and sending codes via email after payment confirmation.	Medium
16	User Activity Monitoring & Reports	Admins can monitor user activities, generate usage reports, and track metrics like daily activity, new registrations, and inactive users.	Low
17	Real-Time Processing Feedback	Displays a progress bar or status indicator during the 3D model generation process, helping manage user expectations on processing times.	Low
18	Guided Model Capture Tutorial	Provides tutorials and tips on capturing optimal photos to enhance the quality of 3D models. Helps improve user experience and model accuracy.	Low
19	Ads Display for Free Users	Displays sponsored ads to free users, generating revenue while offering ad-free experience for premium users.	Low
20	Interactive Tutorials	Provides a help icon for users to access step-by-step guides and tooltips, enhancing the overall user onboarding experience.	Low

5. Non-Functional Requirements

To ensure 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

a) Scalability

- Goal: Ensure the system can handle increased user loads smoothly, even as the user base grows.
- How: Start by optimizing the existing database for better performance. As demand increases, consider cloud-based solutions to scale up without significant rework.

b) Security and Privacy

- **Goal**: Protect user data, especially for uploaded images and 3D models, ensuring data privacy and compliance with regulations.
- **How:**
  - Use HTTPS for secure communication.
  - Implement strong user authentication (e.g., complex passwords, two-factor authentication).
  - Encrypt data at rest and in transit (e.g., using AES-256).
  - Ensure compliance with privacy laws like GDPR.

c) Performance

- **Goal**: Deliver fast and efficient 3D model processing with minimal wait times. Aim for a maximum model generation time of under 60 seconds for standard photo sets.
- **How**:
  - Optimize algorithms for image processing and 3D reconstruction.
  - Leverage parallel processing and GPU acceleration to speed up computations.
  - Implement real-time performance monitoring to quickly detect and fix slowdowns.

d) Usability

- **Goal**: Create an intuitive and user-friendly interface that’s easy for anyone to navigate, even without technical expertise.
- **How**:
  - Follow consistent UI/UX patterns and provide clear instructions with tooltips.
  - Ensure accessibility (e.g., screen reader support, keyboard navigation).
  - Conduct usability testing and gather feedback to refine the user experience.

e) Compatibility

- **Goal**: Ensure the app works well across all popular browsers and devices to reach a broad audience.
- **How**:
  - Test the platform on Chrome, Firefox, Safari, and Edge.
  - Use responsive design frameworks (e.g., Bootstrap) to adapt to various screen sizes, including desktops, tablets, and smartphones.

f) **Availability and Reliability**

- **Goal**: Provide a stable and reliable platform with minimal downtime, aiming for 99.9% uptime.
- **How**:
  - Implement redundancy in servers and databases.
  - Set up automated data backups and a disaster recovery plan.
  - Use monitoring tools (e.g., New Relic, Datadog) to track uptime and system health.

g) Maintainability

- **Goal**: Make the system easy to update and maintain, enabling quick bug fixes and feature enhancements.
- **How**:
  - Use modular code structures and adhere to coding best practices.
  - Set up automated testing and CI/CD pipelines for smooth deployments.
  - Maintain clear and comprehensive documentation for developers.

h) Future Scalability

- **Goal**: Design the system with future growth in mind, allowing for easy addition of new features.
- **How**:
  - Utilize a microservices or service-oriented architecture (SOA) for flexibility.
  - Plan for API integrations to expand capabilities as needed.