VISUAL MENTAL IMAGERY: A VIEW FROM ARTIFICIAL INTELLIGENCE

PRESENTED BY: NICKO A. VENANCIO

OVERVIEW OF COMPUTER VISION IN AI

Computer Vision enables AI to interpret and make decisions based on visual data, mimicking human visual perception. Key capabilities include:

- **Feature Extraction**: Identifying unique attributes within images.
- **Pattern Recognition**: Matching features to stored templates for tasks like object classification and facial recognition.
- **Spatial Interpretation**: Understanding object arrangements and relationships in a scene.
- **Visual Manipulation**: Simulating changes in object positions and perspectives to enhance understanding.

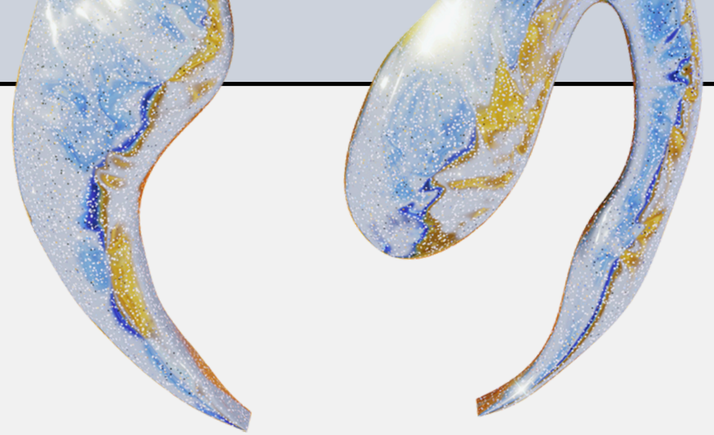
Critical Role of Image Processing

Image Processing is essential in preparing visual data for analysis by AI systems:

- Enhancement: Improves image quality by reducing noise and emphasizing important features.
- Segmentation: Isolates objects or regions within an image for better analysis.
- Feature Extraction: Identifies key characteristics like edges and textures for object classification.
- Transformations: Allows for scaling, rotation, and translation of images to match templates.
- Simulation: Enables AI to predict changes in object appearance and spatial relationships, similar to human visual mental imagery.

Image processing ensures AI systems can accurately analyze and interpret visual data, enabling applications in autonomous vehicles, robotics, and healthcare.

TYPES OF IMAGE PROCESSING TECHNIQUES



1. **Feature Extraction** involves identifying distinct attributes or "features" in an image, like edges, corners, or textures, which serve as identifiable patterns for the AI. These features simplify the image, making it easier to recognize objects or match patterns.

- Example: SIFT (Scale-Invariant Feature Transform) identifies and describes local features in images, remaining robust under scaling and rotation. AKAZE, an alternative to SIFT, is often used in applications needing fast feature extraction.
- Application: Feature extraction is vital in facial recognition, where specific features (e.g., eyes, nose, and mouth structure) help match a person's face to a stored template, or in object detection for autonomous vehicles, identifying other cars and pedestrians in various lighting or weather conditions.

2. Image Segmentation divides an image into distinct regions or objects, isolating the elements that AI needs to analyze. It is foundational for applications requiring object localization or boundary recognition.

- Example: Mask R-CNN (Region-based Convolutional Neural Networks) segments images by generating masks around objects, enabling precise boundary detection.
- Application: Medical imaging relies on segmentation to detect and isolate organs, tissues, or anomalies within MRI or CT scans, assisting radiologists in diagnostics. In autonomous driving, segmentation separates

3. Object Detection

- **Definition:** Identifies and locates objects within images.
- **Example:** YOLO (You Only Look Once) detects multiple objects in real-time.
- **Application:** Security surveillance, detecting objects in real-time video feeds.

CASE STUDY OVERVIEW

It discusses various applications of visual mental imagery in AI, each relying on image processing to simulate human-like reasoning capabilities. For instance, one AI application involves **template-based visual** search, where image processing enables the system to locate a specific visual target within a complex environment. Here, an image template of the target is matched against various sections of an input image to identify the closest match. This method is essential in applications like face recognition, medical imaging, and traffic sign detection, where the system processes visual information to achieve high accuracy in locating and identifying specific patterns.

The challenges addressed by such an application include accurate pattern matching, efficient processing of visual data, and robustness to variations like angle, lighting, or noise. Using image processing, these AI systems must differentiate the target image within diverse backgrounds, requiring the system to be both computationally efficient and adaptable to slight changes in visual input

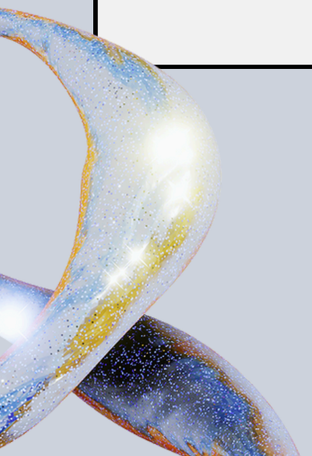
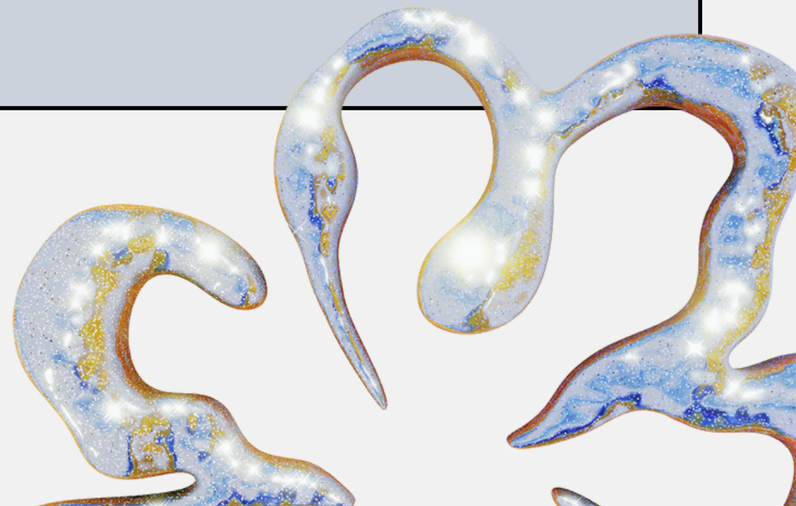


IMAGE PROCESSING IMPLEMENTATION

Model Design

1. **Feature Extraction:** The model begins by analyzing images AAA, BBB, and CCC to identify key visual features such as shape, color, orientation, and size. Instead of using the raw images, the model represents each image as a set of extracted features, creating a simpler, abstract representation.
2. **Feature Comparison**: The model then compares the features of AAA and BBB to find differences or changes (e.g., shape change, color shift, or rotation angle).
3. **Application to CCC:** Once the relationship between AAA and BBB is understood in terms of feature changes, the model applies similar changes to the features of CCC to generate a hypothetical answer.
4. **Answer Selection**: The model compares this generated answer to the candidate images, selecting the one with the closest matching features.

By focusing on feature changes, this model simplifies the problem and helps the AI system identify the underlying relationship in a way that's less reliant on complex image transformations.



CONCLUSION

In conclusion, this feature-based model solves geometric analogy problems by identifying key features in each image and detecting how these features change between pairs. By applying similar changes to the third image, the model generates an answer that best matches the pattern. This approach simplifies the problem, allowing the AI to focus on feature relationships rather than complex image transformations, making it an efficient and effective solution.

THANK YOU, *very much*