

The background is a spiral-bound notebook with a brown cover and silver rings on the left. The page is a light beige color with a fine grid pattern. At the top, the words "GMU MIMICAVE" are written in large, 3D, yellow-to-brown gradient letters. Below the letters, a 3D scene shows a person with a blue backpack climbing a brown ledge. The person is on the right side of the ledge, reaching up. Another person is visible on the ledge below, also climbing. The scene is lit from the left, casting shadows to the right.

GMU MIMICAVE

Juergen Symanzik

Edward J. Wegman

J. Patrick Vandersluis

Qiang Luo

Fernando Camelli

Antoinette Dzubay

Xiadong Fu

Nkem-Amin Khumbah

Rida E. A.Moustafa

Robert L. Wall

Ying Zhu

3rd International Immersive Projection Technology Workshop

May 11, 1999

Outline of Talk

- ✓ The CAVE™ VR Environment
- ✓ Motivation for the MiniCAVE™
- ✓ The MiniCAVE™ Environment
 - Windows NT/Pentium II Implementation
 - Voice Interface
- ✓ Technical Challenges

Trademark Notes

- ✓ CAVE is a trademark of the Electronic Visualization Laboratory at the University of Illinois, Chicago
- ✓ MiniCAVE is a trademark of the Center for Computational Statistics at George Mason University

CAVE Concept

- ✓ A Projection-based Immersive VR System
 - Silicon Graphics-based with 8 to 12 processors
 - RE² or RE Infinity graphics engines
 - CRT-based projection system
 - Stereographics Crystal Eyes shutter glasses
 - Head tracking
 - Usually 3 to 6 wall cube

CAVE Strengths

- ✓ Effective immersive environment
 - Lightweight non-intrusive glasses
 - Can see own hands and other participants
- ✓ Effective for group VR
 - Good tool for group collaboration

CAVE Weaknesses

✓ CRT Projectors

- Projectors not very bright
- Shock, vibration & heat, hard to keep focus
- Geometric distortion at wall interfaces

✓ Tracking

- One user tracked, distorted stereo for users not at viewpoint

✓ User Interface

- Usually 3-D extension of desktop metaphor

CAVE Weaknesses

✓ Expensive

- ~\$1,000,000 fully outfitted
- ~\$600,000 SGI computers
- ~\$30,000 per projector

Motivation

- ✓ Installed MATLAB 5 on SGI Onyx and Pentium
 - Benchmarks on 200 megahertz Pentium Pro (\$3000) and 200 megahertz SGI Onyx (\$120,000) similar
- ✓ Liquid Crystal Projectors sharp, bright, and stable under shock, vibration and temperature variation
- ✓ Stereographics Crystal Eyes technology available for Windows NT

MiniCAVE Concept

- ✓ Windows NT/Intel Pentium II 466 mhz
- ✓ LCD-based projection systems
- ✓ 12 ft cubes scaled to 6 ft cubes
- ✓ Tracking optional, reduced latency
- ✓ Voice command metaphor
- ✓ ~\$100,000 entry level

VR from Workstation to PC

✓ Project Purpose

- Can the NT workstation really match the SGI workstation in 3D graphics area?
- Possibility of VR implementation in PC environment, especially MiniCAVE
- Explore the hardware and software capacities of PC for VR application

A spiral-bound notebook with a brown cover and a cream-colored page. The spiral binding is on the left side. A horizontal line is drawn across the page, separating the header area from the main content.

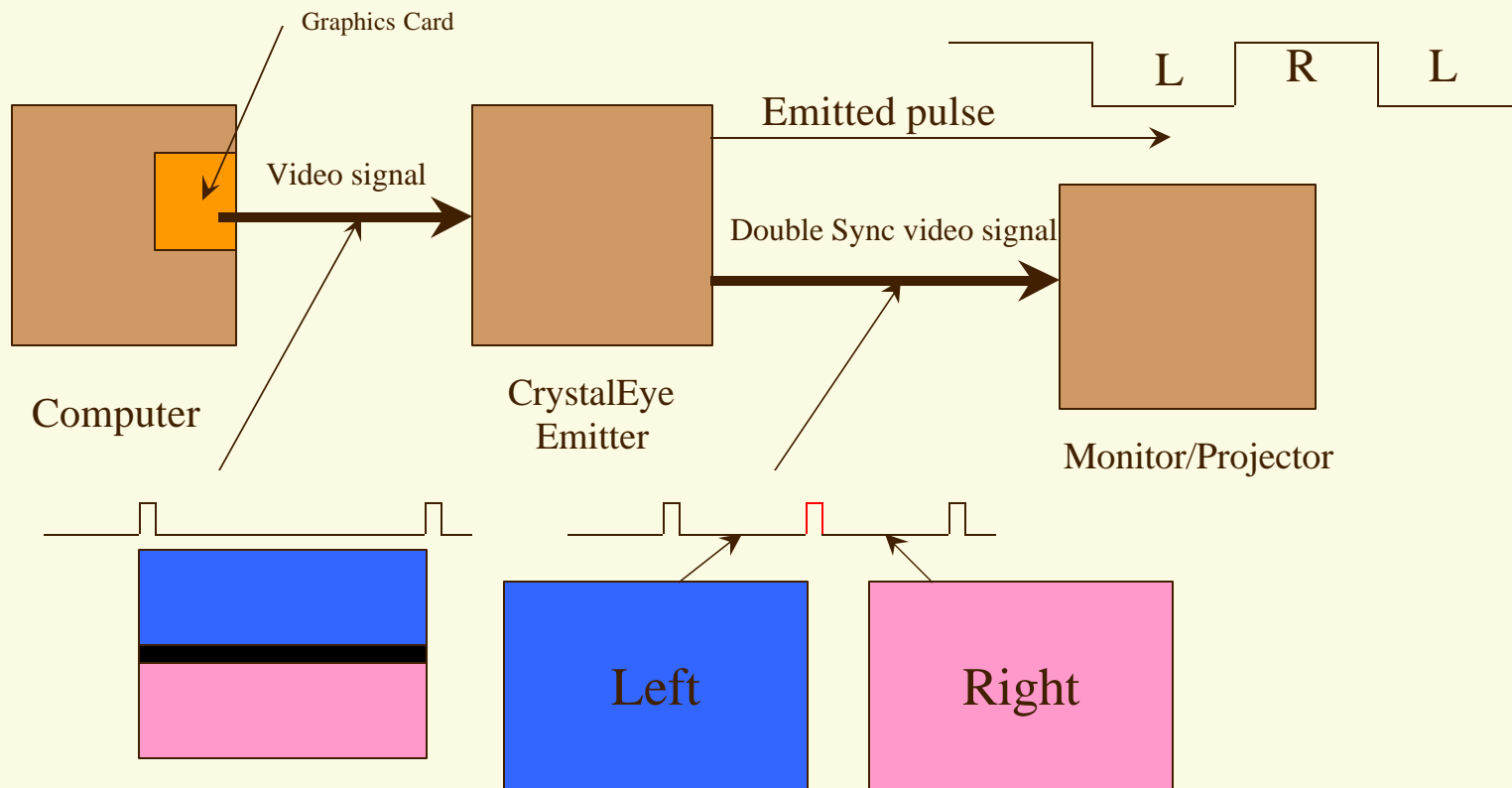
✓ Initial Approach

1. Porting one SGI GL application, SkyFly, to OpenGL, which is a platform-independent 3D API.
2. Porting this application to NT environment.
3. Performance comparison between SGI and NT.
4. Stereo display on PC using CrystalEyes.
5. Controlled by voice command.

✓ Stereo using CrystalEyes

- Above-below stereo
- Image resolution 1024x384 each eye
- Vertical refresh rate 120-150 (60-75 each eye)
 - SGI monitor can handle both 120 and 150
 - CRT projector can only handle 120 refresh rate

✓ Principles of above-below CrystalEyes stereo



Speech Motivation

- ✓ User Interfaces (Van Dam)
- ✓ Shortcuts in XGobi
- ✓ User Controls in the C2 Stats Application

Speech Recognition Products

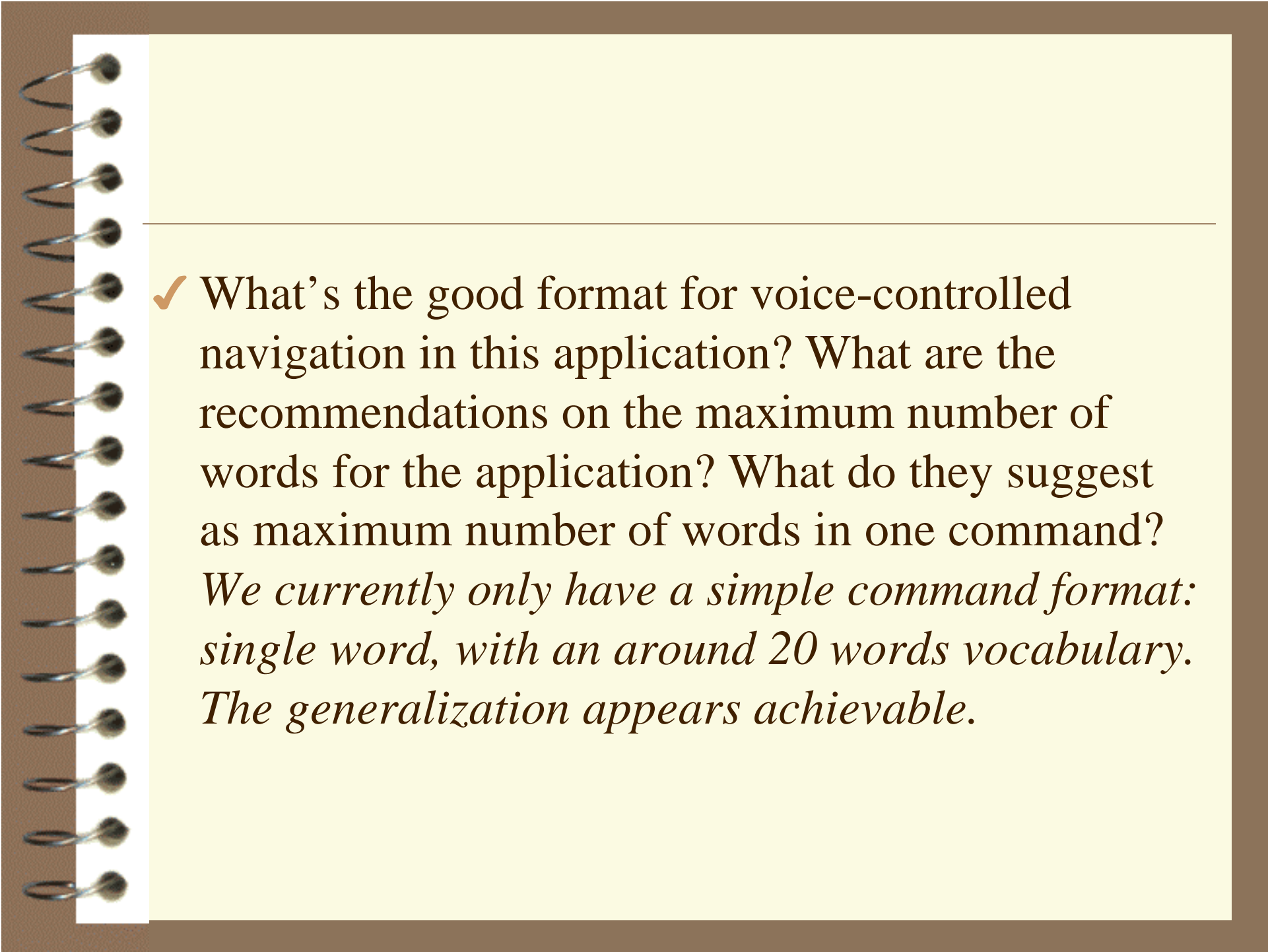
Dragon Dictate

IBM ViaVoice

Speech Recognition Technology

Evaluation

1. Can you train the software to understand additional words?
2. Is the software speaker dependent or speaker independent?
3. Can you store different pronunciations of one word into the same database so a spoken word is compared with different pronunciations?
4. How good is performance (i.e., percentage of correctly identified words before and after training)?
5. Does accuracy depend on speaker / accent / training?
6. What improves accuracy relative to the various sources of imprecision?
7. How long does it take to reach various levels of performance?
8. What if we aren't able to train? What should we expect?

- 
- A graphic of a spiral-bound notebook with a brown cover and a light yellow page. The spiral binding is on the left side. A horizontal line is drawn across the page, separating the top blank area from the text below.
- ✓ What's the good format for voice-controlled navigation in this application? What are the recommendations on the maximum number of words for the application? What do they suggest as maximum number of words in one command?
We currently only have a simple command format: single word, with an around 20 words vocabulary. The generalization appears achievable.

A graphic of a spiral-bound notebook with a brown cover and a cream-colored page. The spiral binding is on the left side. A horizontal line is drawn across the page, separating the header area from the main content area.

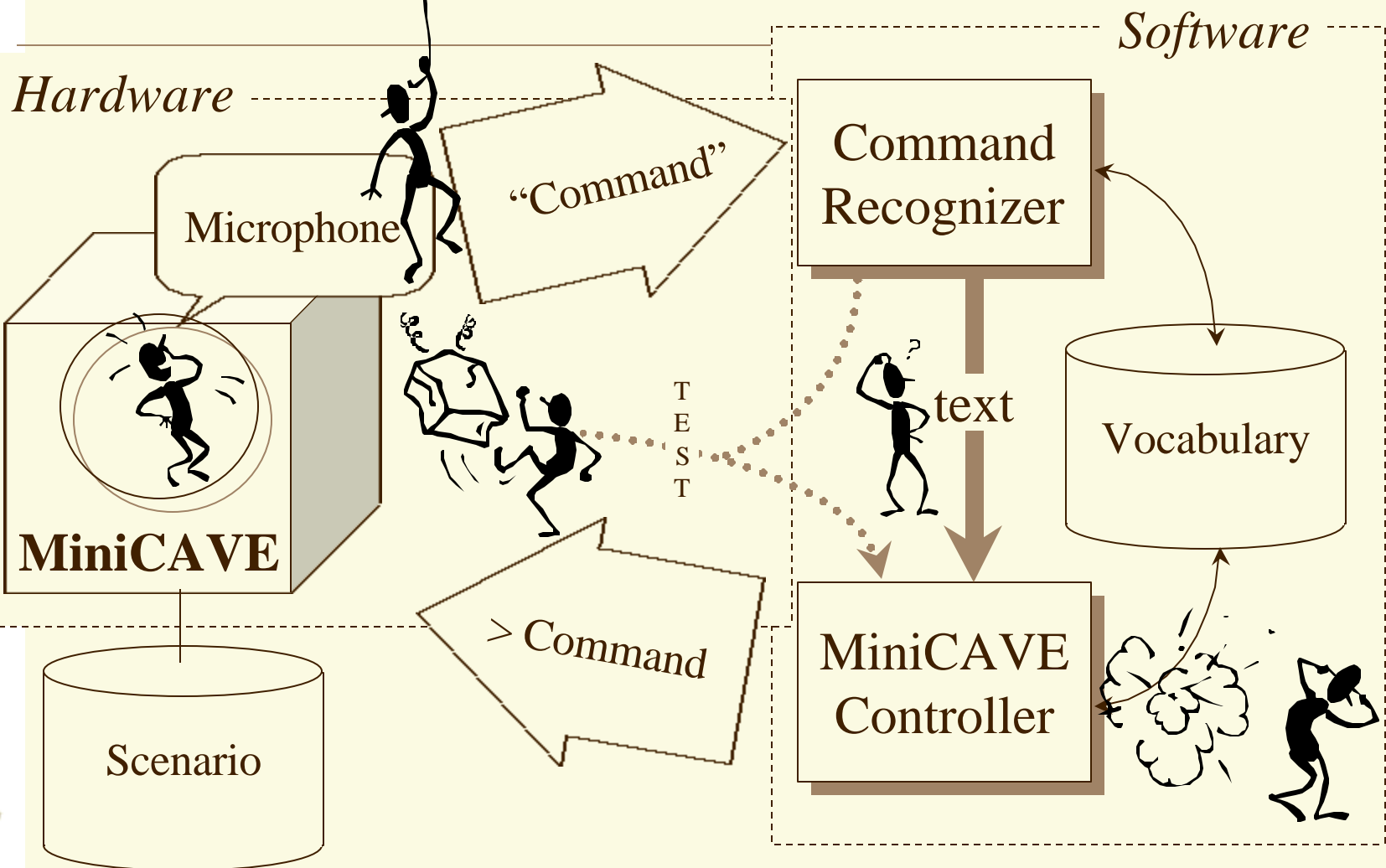
✓ Voice Control

- Command Set

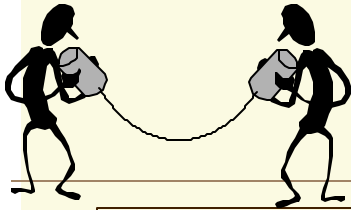
left, right, up, down, fast, slow, forward, reverse, stop, start.

- Link directly to Dragon Dictate (locally), or link to Custom-DLL(network possible)

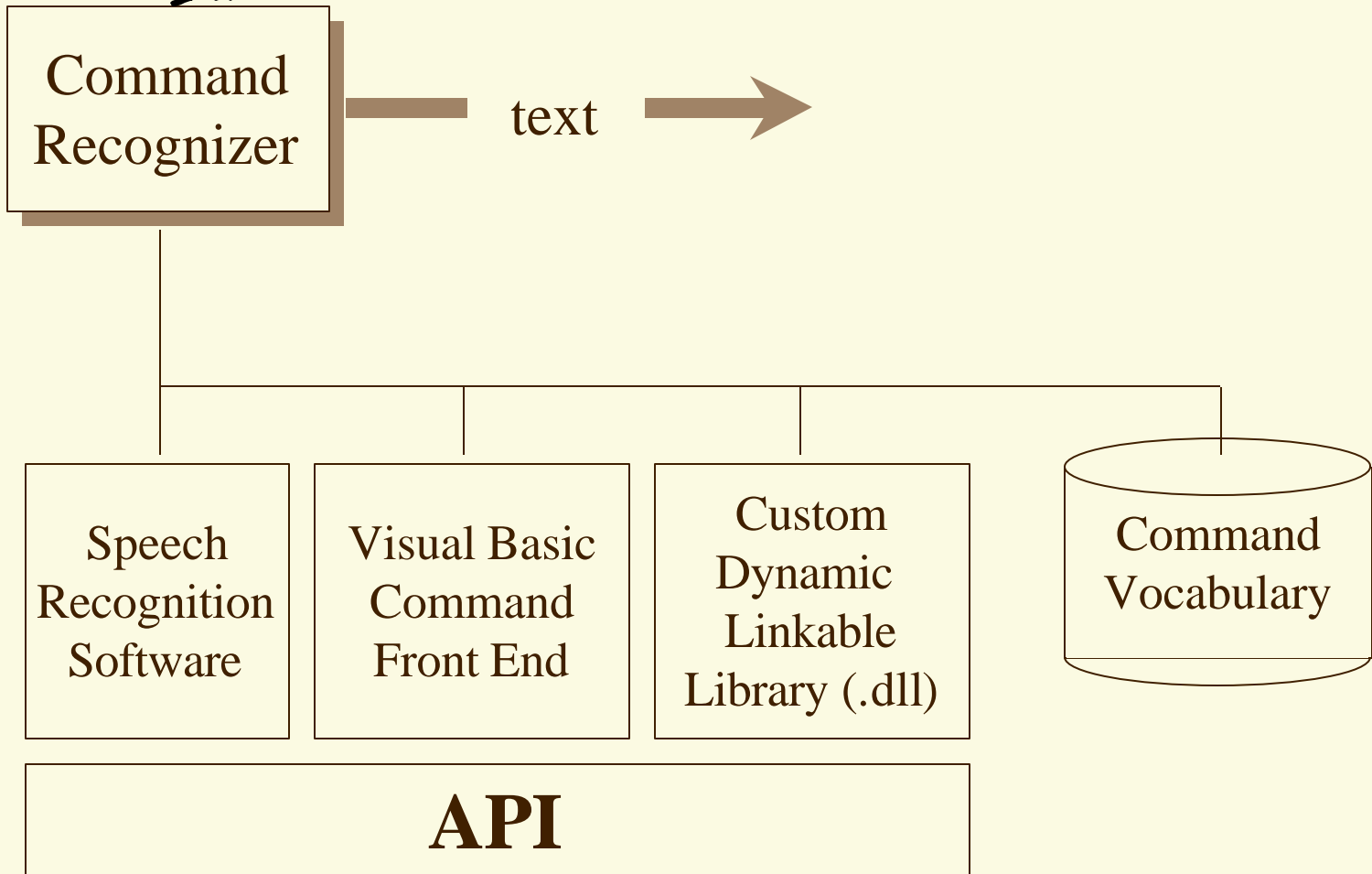
Overview of Voice Interface



Command Recognizer

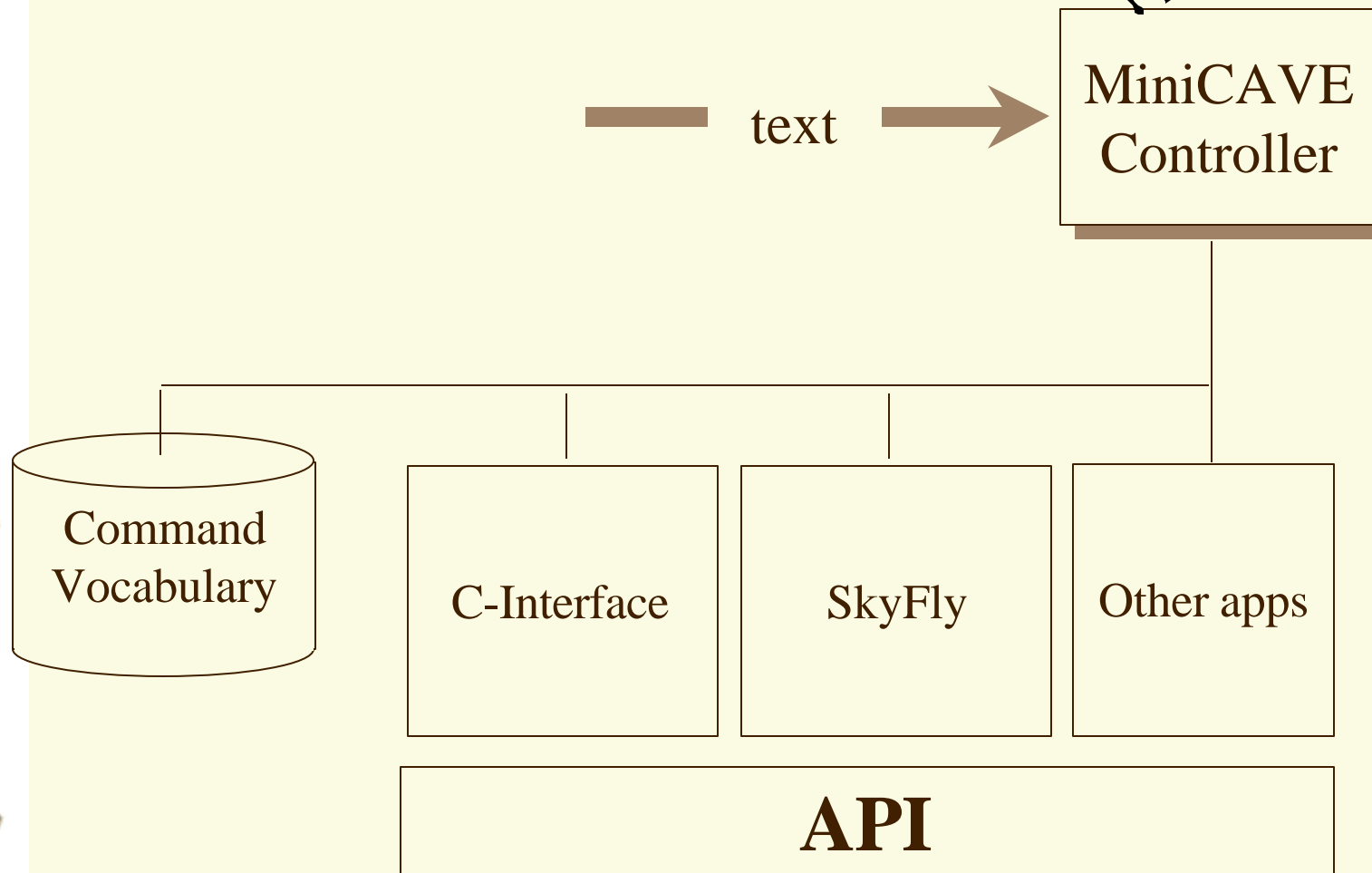
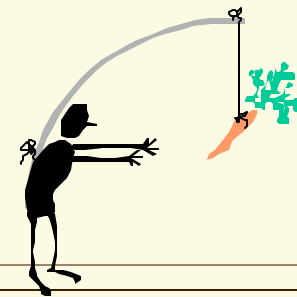


Requirement: Capture Voice Command & Output Text



MiniCAVE Controller

Requirement: Capture & Execute Text Command



Approach

- ✓ Integrate successfully on one machine, then attempt a multi-machine solution
- ✓ Recognition of spoken word causes delay in SkyFly program on 300 mhz Pentium II - much better on 466 mhz Pentium II

Technical Challenges - Successes

- ✓ Port of SkyFly Stereoscopic Demo to NT successful with adequate frame rates on 300 Megahertz Machine
- ✓ CrystalEyes interface on NT successful
- ✓ Voice Recognition using Dragon Dictate successful
 - But requires Training of Speech Recognizer

Future Directions

- ✓ Speech Enhanced ExplorN
- ✓ MRI Controlled through Speech
- ✓ Stat/GIS Application such as ViRGIS
- ✓ Military Applications to Command and Control

Remaining Challenges

✓ MiniCAVE Libraries

- Compatibility to Existing CAVE Libraries

✓ Projection Systems

- Edge Blending with Digital Projectors
- Digital Projectors Themselves
 - Frame Rates
 - Decay
 - Image Lock / Stereo Lock

Current Status

- ✓ Patent Disclosure Filed
- ✓ CRADA signed with U.S. Army (White Sands Missile Range) - Awaiting Funding
- ✓ Planned EDA/Data Mining Application with Voice Interface
- ✓ Major Delay after 300 mhz Machine had been stolen



Contact

✓ Edward Wegman

– ewegman@gmu.edu

✓ Juergen Symanzik

– symanzik@galaxy.gmu.edu